

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model
Run on: November 26, 2003, 12:39:07 ; Search time 6537 Seconds
(without alignments)
10964.306 Million cell updates/sec

Title: US-10-027-859-1
Perfect score: 1752
Sequence: 1 ttgttttaacaacatgtt.....aaaaaaaaaaaaaaaaaaaa 1752

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0
Searched: 2888711 seqs, 2045481386 residues
Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :	GenEmbl.:
1: gb_ba.:	U91679 Rattus norv
2: gb_htg.:	AC132020 Rattus no
3: gb_in.:	AC127107 Rattus no
4: gb_om.:	AC095777 Rattus no
5: gb_ov.:	AY049085 Mus muscu
6: gb_pat.:	AY049086 Mus muscu
7: gb_ph.:	AC138600 Mus muscu
8: gb_pl.:	AC104542 Mus muscu
9: gb_pr.:	AC139023 Mus muscu
10: gb_ro.:	AF336381 Mus muscu
11: gb_sts.:	BC023511 Homo sapi
12: gb_sy.:	Y08976 H. sapiens m
13: gb_un.:	AC097468 Homo sapi
14: gb_vi.:	Y08979 H. sapiens F
15: em_ba.:	AJ338119 Homo sapi
16: em_fun.:	AJ339729 Homo sapi
17: em_hum.:	AJ339670 Homo sapi
18: em_in.:	AJ340528 Homo sapi
19: em_mu.:	AJ338339 Homo sapi
20: em_om.:	AJ331239 Homo sapi
21: em_or.:	AJ339879 Homo sapi
22: em_ov.:	AJ325918 Homo sapi
23: em_pat.:	AJ325918 Homo sapi
24: em_ph.:	AR232559 Sequence
25: em_pl.:	AR232559 Sequence
26: em_ro.:	BD137836 Hypoxia-r
27: em_sts.:	AF334162 Rattus no
28: em_un.:	AJ339666 Homo sapi
29: em_vi.:	AJ339105 Homo sapi
30: gb_ro.:	AJ322575 Homo sapi
31: em_htg_hum.:	AJ339694 Homo sapi
32: em_htg_inv.:	AJ249590 Danio rer
33: em_htg_other.:	AJ333530 Homo sapi
34: em_htg_mus.:	AJ339388 Homo sapi
35: em_htg_pln.:	AF177538 Danio rer
36: em_htg_rod.:	AP003148 Mus muscu
37: em_htg_mam.:	S66169 Erg-3-immun
38: em_sy.:	AB073078 Mus muscu
39: em_htgo_hum.:	AB073079 Mus muscu
40: em_htgo_mus.:	AB073080 Mus muscu
41: em_htgo_other.:	S72621 EWS...erg {

ALIGNMENTS

RESULT 1	RNU91679	1752 bp	mrna	linear	ROD 07-APR-1998
RNU91679	Rattus norvegicus	ETS domain	transcription factor	Pet-1	mrna,
LOCUS	complete cds.				
DEFINITION	U91679				
ACCESSION	U91679.1	GI:3033418			
VERSION					
KEYWORDS					
SOURCE					
ORGANISM					
	Rattus norvegicus (Norway rat)				
	Rattus norvegicus				
	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;				
	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;				
	Rattus.				
REFERENCE	1 (bases 1 to 1752)				
AUTHORS	Fyodorov,D., Nelson,T. and Deneris,E.				

Pred. No. is the number of results predicted by chance to have a

TITLE Pet-1, a novel ETS domain factor that can activate neuronal nachr
JOURNAL J. Neurobiol. 34 (2), 151-163 (1998)
MEDLINE 98127904
PUBMED 9468386
REFERENCE 2 (bases 1 to 1752)
AUTHORS Eydorov,D., Nelson,T. and Deneris,E.
TITLE Direct Submission
JOURNAL Submitted (01-MAR-1997) Neurosciences, Case Western Reserve
University, 2109 Adelbert Rd., Cleveland, OH 44106, USA
FEATURES
source Location/Qualifiers
1..1752
/organism="Rattus norvegicus"
/mol_type="mRNA"
/db_xref="taxon:10116"
/cell_line="PC12"
/tissue_type="pheochromocytoma"
112..1134
/codon_start=1
/product="ETS domain transcription factor PET-1"
/protein_id="AAC12859.1"
/db_xref="GI:3033419"
/translation="MEDPFGGALGERVPAHPPOHPLTAHSSSTPAPGWMQLODP
LPHTLAAASRQALPPAASPATLPCHPSPRAGITPSAKLTCPVRSPPSPTAQSPA
AMROSGTSQPLLIINNYLPDPVGDGLFKEGKSPSMGLSPAVOKGSGOIQLWQFLLEL
ADRANAGCIAWEGHGFELKLPDPEVARWGERKSKPMNVYDKLSRALRYVYDKNIMS
KVHGKRYARDFDGLAQACOPPPAHAAHAAAAAQAQDGLYKLPAGLAPLPFP
GLSKLNLMAASAGVAPAGFSYWPGFNATAAAATAALYPTFGLPQPPFGVAANAAS
LGHHYH"

CDS
polyA_signal 1705..1710
BASE COUNT 356 a 613 c 412 g 371 t
ORIGIN

Query Match 100.0%; Score 1752; DB 10; Length 1752;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1752; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TTTGTTTAAACAAATGTTTATTAGAAAAGTAAATAATTGCATAGTCTTAGTACTTG 60
Db 1 TTTGTTTAAACAAATGTTTATTAGAAAAGTAAATAATTGCATAGTCTTAGTACTTG 60
Qy 61 AACATCAAGTGATTTCAATGAACCGTGAGTATCTTTCAATGTAACAGTTCTAGATGGAAGAC 120
Db 61 AACATCAAGTGATTTCAATGAACCGTGAGTATCTTTCAATGTAACAGTTCTAGATGGAAGAC 120
Qy 121 CCAGTGCGCTCTCTGGGGAGAGGTTCCAGCCCCCAGCCCCCTCAGCCCCATCCC 180
Db 121 CCAGTGCGCTCTCTGGGGAGAGGTTCCAGCCCCCAGCCCCCTCAGCCCCATCCC 180
Qy 181 CTCACAGTCACTCTCCAGTACACCGGCACCGGATGGGCTGGGATGCAGCTCCAGGAC 240
Db 181 CTCACAGTCACTCTCCAGTACACCGGCACCGGATGGGCTGGGATGCAGCTCCAGGAC 240
Qy 241 CCCCTCCCTCTCACACACCTCGGCTCCCGCCAGGCTTTGCCGACCCCGCG 300
Db 241 CCCCTCCCTCTCACACACCTCGGCTCCCGCCAGGCTTTGCCGACCCCGCG 300
Qy 301 GCGTCTACTCTTCCGTACCCACAGTCACACGGGGGTATCGGACCCCAAGGCA 360
Db 301 GCGTCTACTCTTCCGTACCCACAGTCACACGGGGGTATCGGACCCCAAGGCA 360
Qy 361 AAGCTGAGTGCCTCCCGCTGCGGTCCCGCCCATCTCCACCGCCAGTCCCGGAGCG 420
Db 361 AAGCTGAGTGCCTCCCGCTGCGGTCCCGCCCATCTCCACCGCCAGTCCCGGAGCG 420
Qy 421 ATGAGACAGCGGACCTCCAGCGCTGTGTATCAACATGTACCTACAGATCCCGTC 480
Db 421 ATGAGACAGCGGACCTCCAGCGCTGTGTATCAACATGTACCTACAGATCCCGTC 480
Qy 481 GGAGATGCTCTTTTAAAGGAAGGAAGAGCCGAGCTGGGGCGCTGAGCCCTGCGGTA 540
Db 481 GGAGATGCTCTTTTAAAGGAAGGAAGAGCCGAGCTGGGGCGCTGAGCCCTGCGGTA 540

Qy 541 CAGAAAGGACGCGGAGATCCAGTTGTGGCAGTTTCTACTGGAGCTGTCTGGCAGACCGC 600
Db 541 CAGAAAGGACGCGGAGATCCAGTTGTGGCAGTTTCTACTGGAGCTGTCTGGCAGACCGC 600
Qy 601 GCGAAACGCGGCTGCATCGCGTGGAGGCGGCGCACCGCGAGTTCAAGCTCACGACGCC 660
Db 601 GCGAAACGCGGCTGCATCGCGTGGAGGCGGCGCACCGCGAGTTCAAGCTCACGACGCC 660
Qy 661 GACGAGTGGCGGACGCTGGGGCGAGCGCAAGAGAGCCCAATATATGAACCTACGACAAG 720
Db 661 GACGAGTGGCGGACGCTGGGGCGAGCGCAAGAGAGCCCAATATATGAACCTACGACAAG 720
Qy 721 CTAAGTCGAGCAGTCCGCTACTACTACACAAAACATCATGAGCAAGGTCACGGCAAG 780
Db 721 CTAAGTCGAGCAGTCCGCTACTACTACACAAAACATCATGAGCAAGGTCACGGCAAG 780
Qy 781 CGCTACGCTTACCGCTTTGACTTCCAGGGCTGGCAGAGGTTGCCAGCCACACCCGCG 840
Db 781 CGCTACGCTTACCGCTTTGACTTCCAGGGCTGGCAGAGGTTGCCAGCCACACCCGCG 840
Qy 841 CAGCCCAACGCGCGCTGCGCGCGCAGCGCAGCGCGCGCCGCGCAGGATGGCGCACTT 900
Db 841 CAGCCCAACGCGCGCTGCGCGCGCAGCGCAGCGCAGCGCGCGCCGCGCAGGATGGCGCACTT 900
Qy 901 TACAAGCTCCCGCTGGCTTGGCTCCACTGCCCTTCCCGCGCCTCTCCAAACTCAACCTT 960
Db 901 TACAAGCTCCCGCTGGCTTGGCTCCACTGCCCTTCCCGCGCCTCTCCAAACTCAACCTT 960
Qy 961 ATGCGAGCTCCGCGCGCTGGCGCGCGCTGGCTTCTTACTGGCTGTGTCCTCAACGCC 1020
Db 961 ATGCGAGCTCCGCGCGCTGGCGCGCGCTGGCTTCTTACTGGCTGTGTCCTCAACGCC 1020
Qy 1021 ACCGCGCTCGCGCGCAGCGTGGCTCTTACCCAAACCGCGGCTTGCAGCCCCCTCCC 1080
Db 1021 ACCGCGCTCGCGCGCAGCGTGGCTCTTACCCAAACCGCGGCTTGCAGCCCCCTCCC 1080
Qy 1081 GGGCCCTTTGGCGGCTGGCGCGCTTCCGACTTGGGGGTCAATATCACTAGACGGGA 1140
Db 1081 GGGCCCTTTGGCGGCTGGCGCGCTTCCGACTTGGGGGTCAATATCACTAGACGGGA 1140
Qy 1141 CGGCGGCTCGAGTGGGGCTCTCCACACAGCGCAGTGACCAATCCCATCTCATCTGG 1200
Db 1141 CGGCGGCTCGAGTGGGGCTCTCTCCACACAGCGCAGTGACCAATCCCATCTCATCTGG 1200
Qy 1201 GAGAGCCCCGAGATTTCCCGAGCGTCTTTTACCACAGATTTCTGTGACGAGCGCT 1260
Db 1201 GAGAGCCCCGAGATTTCCCGAGCGTCTTTTACCACAGATTTCTGTGACGAGCGCT 1260
Qy 1261 CCCAGCCGAGGAAGAGATGGGAGCCTCTGAGGTCTTCTTGAATACGAGGCTTCC 1320
Db 1261 CCCAGCCGAGGAAGAGATGGGAGCCTCTGAGGTCTTCTTGAATACGAGGCTTCC 1320
Qy 1321 AGGCTCCCATATATCATCACCCAGGAAGGTGATGTGCTCCCACTTTTAAATTTTCTCT 1380
Db 1321 AGGCTCCCATATATCATCACCCAGGAAGGTGATGTGCTCCCACTTTTAAATTTTCTCT 1380
Qy 1381 CCAAGTCTCAGATTTGGAATCTCCCGTCTTTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1440
Db 1381 CCAAGTCTCAGATTTGGAATCTCCCGTCTTTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1440
Qy 1441 CTTCTCTTTATGACCCCTAGTTTCTGTTTGTGTTTTTTTTTTTTTCTCTCTCTCTCTC 1500
Db 1441 CTTCTCTTTATGACCCCTAGTTTCTGTTTGTGTTTTTTTTTTTTTCTCTCTCTCTCTC 1500
Qy 1501 ATTTTTTTCTCTCCACGACCTACTCCAAACGGTAGTACTCGGTAGTACTCTCGAGGCT 1560
Db 1501 ATTTTTTTCTCTCCACGACCTACTCCAAACGGTAGTACTCGGTAGTACTCTCGAGGCT 1560
Qy 1561 TCTCACACTCCCTTTTGGGATATGAGAGCATCAAAAAACATCTCTGTGTTGTCTCATC 1620
Db 1561 TCTCACACTCCCTTTTGGGATATGAGAGCATCAAAAAACATCTCTGTGTTGTCTCATC 1620
Qy 1621 CTTATCCCAACACTCTGGCTTCCCTTCCATACCACACTCTGGGCCCAGGACCTCTG 1680

```

Db      1621  CCTATCCCAACACTCTGGCTTCGCTCCCTTCCATACCACACTCTGGCCCAAGGACCTCG 1680
|||||
QY      1681  TCTGTATATATCTCTTCAGCCCATTAAGATCCCAAGCTTCAAAAAAAAAAAAAAAAAAAAA 1740
|||||
Db      1681  TCTGTATATATCTCTTCAGCCCATTAAGATCCCAAGCTTCAAAAAAAAAAAAAAAAAAAAA 1740
|||||
QY      1741  AAAAAAAAAAAAAA 1752
|||||
Db      1741  AAAAAAAAAAAAAA 1752
|||||

RESULT 2
AC132020
LOCUS      Rattus norvegicus clone CH230-250012, WORKING DRAFT SEQUENCE.
DEFINITION AC132020      205178 bp      DNA      linear      HTG 15-NOV-2002
ACCESSION AC132020
VERSION    HTG; HTGS_PHASE2; HTGS_DRAFT; HTGS_FULLTOP.
KEYWORDS   Rattus norvegicus (Norway rat)
SOURCE     Rattus norvegicus
ORGANISM   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
1 (bases 1 to 205178)
Muzny,D.Marie., Metzker,M.Lee., Abramson,S., Adams,C., Alder,J.,
Allen,C., Allen,H., Alsbrooks,S., Amin,A., Anguiano,D.,
Anyalebechi,V., Aoyagi,A., Ayodeji,M., Baca,E., Baden,H.,
Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F.,
Biswal,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M.,
Bryant,N., Buhay,C., Burch,P., Burrell,K., Calderon,E.,
Cardenas,V., Carter,K., Cavazos,I., Ceasar,H., Center,A.,
Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J.,
Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L.,
Davila,M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D.,
Delgado,O., Denison,S., Deramo,C., Ding,Y., Dinh,H., Divya,K.,
Draper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Eaves,K.,
Egan,A., Escotto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G.,
Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P.,
Fraser,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Garza,M.,
Gebregeorgis,E., Geer,K., Gill,R., Grady,M., Guerra,W., Guevara,W.,
Gunaratne,P., Haaland,W., Hamil,C., Hamilton,C., Hamilton,J.,
Harvey,Y., Havlak,P., Haves,A., Henderson,N., Hernandez,J.,
Hernandez,R., Hines,S., Hladun,S.L., Hodgson,A., Hogues,M.,
Hollins,B., Howells,S., Hulyk,S., Hume,J., Idlebird,D., Jackson,A.,
Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivet,A.,
Karpathy,S., Kelly,S., Kelly,S., Khan,Z., King,L., Kovar,C.,
Kowis,C., Kraft,C.L., Lebow,H., Levan,J., Lewis,D., Li,Z., Liu,J.,
Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J.,
Lorensuhewa,L., Loulsegged,H., Lozado,R.J., Lu,X., Ma,J.,
Maheshwari,M., Mahindartne,M., Mahmoud,M., Malloy,K., Mangum,A.,
Mangum,B., Mapua,P., Martin,K., Martin,R., Martinez,E.,
Mawhney,S., McLeod,M.P., McNeill,T.Z., Meenen,E.,
Milosavljevic,A., Minter,G., Minja,E., Montemayor,J., Moore,S.,
Morgan,M., Morris,K., Morris,S., Munidasa,M., Murphy,M., Nair,L.,
Nankervis,C., Neal,D., Newton,N., Nguyen,N., Norris,S.,
Nwaokemelehen,O., Okwuonu,G., Olarnpusagoon,A., Pal,S., Parks,K.,
Pasternak,S., Paul,H., Perez,A., Perez,L., Pfannkuch,C.,
Plopper,F., Polindexter,A., Popovic,D., Primus,E., Pu,L.,
Puafo,M., Quiroz,J., Rachlin,E., Reeves,K., Regier,M.A., Reigh,R.,
Reilly,B., Reilly,M., Ren,Y., Reuter,M., Richards,S., Riggs,F.,
Rives,C., Rodkey,T., Rojas,A., Rose,M., Rose,R., Ruiz,S.J.,
Sanders,W., Savery,G., Scherer,S., Scott,G., Shatsman,S., Shen,H.,
Shetty,J., Shvartsbeyn,A., Sisson,I., Sitter,C.D., Smajd,D.,
Sneed,A., Sodergren,E., Song,X.-Z., Sorelle,R., Soes,J.,
Steinle,M., Strong,R., Sutton,A., Svatek,A., Tabor,P., Taylor,C.,
Taylor,T., Thomas,R., Thomas,S., Tingey,A., Trejos,Z., Usmani,K.,
Valas,R., Vera,V., Villasana,D., Waldron,L., Walker,B., Wang,J.,
Wang,Q., Warren,J., Warren,R., Warren,R., Wei,X., White,F.,
Williams,G., Willson,R., Wleczyk,R., Wooden,H., Worley,K.,
Wright,D., Wright,R., Wu,J., Yakub,S., Yen,J., Yoon,L., Yoon,V.,
Yu,F., Zhang,J., Zhou,J., Zhou,X., Zhao,S., Dunn,D., von
Niederhausern,A., Weiss,R., Smith,D.R., Holt,R.A., Smith,H.O.,

```

```

TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
FEATURES
source
BASE COUNT
ORIGIN
Query Match
Best Local Similarity
Matches
QY
Db

```

```

Weinstock,G. and Gibbs,R.A.
Direct Submission
Unpublished
2 (bases 1 to 205178)
Rat Genome Sequencing Consortium.
Direct Submission
Submitted (29-AUG-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 205178)
Rat Genome Sequencing Consortium.
Direct Submission
Submitted (15-NOV-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
The sequence in this assembly is a combination of BAC based reads
and whole genome shotgun sequencing reads assembled using Atlas
(http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described
in the feature table below represents a scaffold in the Atlas
assembly (a 'contig-scaffold'). Within each contig-scaffold,
individual sequence contigs are ordered and oriented, and separated
by sized gaps filled with Ns to the estimated size. The sequence
may extend beyond the ends of the clone and there may be sequence
contigs within a contig-scaffold that consist entirely of whole
genome shotgun sequence reads. Both end sequences and whole genome
shotgun sequence only contigs will be indicated in the feature
table.
----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: KBPD
Center clone name: CH230-250012
----- Summary Statistics
Assembly program: Phrap; version 0.990329
Consensus quality: 197555 bases at least Q40
Consensus quality: 198713 bases at least Q30
Consensus quality: 199469 bases at least Q20
Estimated insert size: 203019; sum-of-contigs estimation
Quality coverage: 8x in Q20 bases; sum-of-contigs estimation
-----
* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 1 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* provided by the submittor.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
*
* 1 205178: contig of 205178 bp in length.
  Location/Qualifiers
    1. 205178
      /organism="Rattus norvegicus"
      /mol_type="genomic DNA"
      /db_xref="taxon:10116"
      /clone="CH230-250012"
BASE COUNT 51577 a 49768 c 48276 g 50223 t 5334 others
ORIGIN
Query Match 67.2%; Score 1177.6; DB 2; Length 205178;
Best Local Similarity 99.7%; Pred. No. 5e-227;
Matches 1180; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 539 TACCAAGAGCGACGGCGGACATCCAGTTGTGGCAGTTTCTACTGGAGCTCTGGCAGACC 598
Db 48558 TCCCAACAGGACGCGGCGGACATCCAGTTGTGGCAGTTTCTACTGGAGCTCTGGCAGACC 48617

```

Qy	599	CGCGAAGCGCGGCTGCATCGCTGGAGGGCGGCCACGGCGAGTTCAAGCTTCAACGACC	658
Db	48618	CGCGAAGCGCGGCTGCATCGCTGGAGGGCGGCCACGGCGAGTTCAAGCTTCAACGACC	48677
		CGCGAAGCGCGGCTGCATCGCTGGAGGGCGGCCACGGCGAGTTCAAGCTTCAACGACC	48677
Qy	659	CCGACGAGGTGGCGCAGCGCTGGGGCGAGCGCAAGAGCAAGCCCAATATGAATACGACA	718
Db	48678	CCGACGAGGTGGCGCAGCGCTGGGGCGAGCGCAAGAGCAAGCCCAATATGAATACGACA	48737
Qy	719	AGCTAAGTTCGAGCACTCGGCTACTACTACGACAAAAACATCATGAGCAAGGTGCACGCA	778
Db	48738	AGCTAAGTTCGAGCACTCGGCTACTACTACGACAAAAACATCATGAGCAAGGTGCACGCA	48797
Qy	779	AGCGCTAGCCTACCGCTTTGACTTCCAGGGCCTGGCAGAGGCTGCCAGCCACACCG	838
Db	48798	AGCGCTAGCCTACCGCTTTGACTTCCAGGGCCTGGCAGAGGCTGCCAGCCACACCG	48857
Qy	839	CGCAGCGCCACGCGCGCGCTGCCGCGCGCCAGCGCAGCGCGCGCCAGGATGGCGCAC	898
Db	48858	CGCAGCGCCACGCGCGCGCTGCCGCGCGCCAGCGCAGCGCGCGCCAGGATGGCGCAC	48917
Qy	899	TTTACAAGCTCCCGGCTGGTCTGGCTCCACTGGCCCTTCCCGGCGCTCTCAAACCTCAACC	958
Db	48918	TTTACAAGCTCCCGGCTGGTCTGGCTCCACTGGCCCTTCCCGGCGCTCTCAAACCTCAACC	48977
Qy	959	TTATGGCAGGCTCGCGCGCGGTGGCGCGCGCTGGCTTCTTACTTGGCCTGGTCCCAACG	1018
Db	48978	TTATGGCAGGCTCGCGCGCGGTGGCGCGCGCTGGCTTCTTACTTGGCCTGGTCCCAACG	49037
Qy	1019	CCACGCGGCTGGCGCGCGCACGCTCGGCTCTACCCAAACCCGGGCTTGCAGCGCCCTC	1078
Db	49038	CCACGCGGCTGGCGCGCGCACGCTCGGCTCTACCCAAACCCGGGCTTGCAGCGCCCTC	49097
Qy	1079	CCGGGCGCTTTGGCGCGGTGGCGCGCGCTTCGCACATTGGGGGTCAATATCACTAGACGG	1138
Db	49098	CCGGGCGCTTTGGCGCGGTGGCGCGCGCTTCGCACATTGGGGGTCAATATCACTAGACGG	49157
Qy	1139	GACGGCGGGTGCAGTGGGGCTCTCCACACAGCCAGTGCACAAATCCCATCTCATCCT	1198
Db	49158	GACGGCGGGTGCAGTGGGGCTCTCCACACAGCCAGTGCACAAATCCCATCTCATCCT	49217
Qy	1199	GGGAGGAGCCCCGAAGATTTCCCGAGGTTCTTTACACAGATTCGTTGCAGACGG	1258
Db	49218	GGGAGGAGCCCCGAAGATTTCCCGAGGTTCTTTACACAGATTCGTTGCAGACGG	49277
Qy	1259	CTCCAGCCAGGAAGAAAGGATGGGAAGCTCTGAGGTCTTCCTTGAATACGAGGCTT	1318
Db	49278	CTCCAGCCAGGAAGAAAGGATGGGAAGCTCTGAGGTCTTCCTTGAATACGAGGCTT	49337
Qy	1319	CGAGGCTCCCATTTATCATCACCCAGGAGGTGCAATGCTCCCACTTTAATTTTCTC	1378
Db	49338	CGAGGCTCCCATTTATCATCACCCAGGAGGTGCAATGCTCCCACTTTAATTTTCTC	49397
Qy	1379	TTCCAAGTCTCCAGATTTCTGAACTCCCGTCTTTTTTTTCTCTCTCACTGGAGCCCT	1438
Db	49398	TTCCAAGTCTCCAGATTTCTGAACTCCCGTCTTTTTTTTCTCTCTCACTGGAGCCCT	49457
Qy	1439	GCCTTCCTTTTATGACCCCTAGTTTTCTGTTTTGTTTTTTTTTTTTTCTCTCTCTCC	1498
Db	49458	GCCTTCCTTTTATGACCCCTAGTTTTCTGTTTTGTTTTTTTTTTTTTCTCTCTCTCC	49517
Qy	1499	TCATTTTTTTTCTCTCCAGACCTACTCCAAACGGTAGTACTCGGTAGTACCTCGAGG	1558
Db	49518	TCATTTTTTTTCTCTCCAGACCTACTCCAAACGGTAGTACTCGGTAGTACCTCGAGG	49577
Qy	1559	CTTCTCACACTCCCTTTTTCGGGATATGAGAAGCATCAAAAACATCTCTGCTGTGTCCA	1618
Db	49578	CTTCTCACACTCCCTTTTTCGGGATATGAGAAGCATCAAAAACATCTCTGCTGTGTCCA	49637
Qy	1619	TCCTTATCCCAACACTCTGGCTTCGCTCCCTTCCATACCACTCTGGCCCAAGACCT	1678
Db	49638	TCCTTATCCCAACACTCTGGCTTCGCTCCCTTCCATACCACTCTGGCCCAAGACCT	49697
Qy	1679	CGTCTGTATATATTCCTTTTTCAGCCCATTAAGATCCAAAGTTC	1722

REFERENCE
 AUTHORS
 TITLE
 JOURNAL

COMMENT

Baylor Plaza, Houston, TX 77030, USA
 3 (bases 1 to 228187)
 Rat Genome Sequencing Consortium.
 Direct Submission
 Submitted (09-NOV-2002) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA
 On Nov 9, 2002 this sequence version replaced gi:23264396.
 The sequence in this assembly is a combination of BAC based reads
 and whole genome shotgun sequencing reads assembled using Atlas
 (http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described
 in the feature table below represents a scaffold in the Atlas
 assembly (a 'contig-scaffold'). Within each contig-scaffold,
 individual sequence contigs are ordered and oriented, and separated
 by sized gaps filled with Ns to the estimated size. The sequence
 may extend beyond the ends of the clone and there may be sequence
 contigs within a contig-scaffold that consist entirely of whole
 genome shotgun sequence reads. Both end sequences and whole genome
 shotgun sequence only contigs will be indicated in the feature
 table.

----- Genome Center
 Center: Baylor College of Medicine
 Center code: BCM
 Web site: http://www.hgsc.bcm.tmc.edu/
 Contact: hgsc-help@bcm.tmc.edu
 ----- Project Information
 ----- Project name: GNHC
 Center clone name: CH230-206A13
 ----- Summary Statistics
 Assembly program: Phrap; version 0.990329
 Consensus quality: 216089 bases at least Q40
 Consensus quality: 218930 bases at least Q30
 Consensus quality: 220275 bases at least Q20
 Estimated insert size: 220997; sum-of-contigs estimation
 Quality coverage: 8x in Q20 bases; sum-of-contigs estimation

 * NOTE: Estimated insert size may differ from sequence length
 * (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 1 contigs. Gaps between the contigs
 * are represented as runs of N. The order of the pieces
 * is believed to be correct as given, however the sizes
 * of the gaps between them are based on estimates that have
 * provided by the submittor.
 * This sequence will be replaced
 * by the finished sequence as soon as it is available and
 * the accession number will be preserved.
 * 1 228187; contig of 228187 bp in length.

FEATURES
 source
 1..228187
 /organism="Rattus norvegicus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10116"
 /clone="CH230-206A13"
 1..1483
 /note="wgs_end_extension
 clone_end:T7"
 1534..2945
 /note="wgs_end_extension
 clone_end:T7"
 2996..5407
 /note="wgs_end_extension
 clone_end:T7"
 complement(6057..6774)
 /note="clone_boundary
 clone_end:T7
 site:EcoRI
 end_sequence:BH283274"
 complement(225072..225776)
 /note="clone_boundary
 clone_end:Sp6
 site:EcoRI
 end_sequence:BH283275"

BASE COUNT 59820 a 53444 c 52556 g 55801 t 6566 others

ORIGIN

Query Match 67.2%; Score 1177.6; DB 2; Length 228187;
 Best Local Similarity 99.7%; Pred. No. 5.1e-227;
 Matches 1180; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 539 TACAAAGGAGGCGGGCAGATCCAGTTGTGGCAGTTTCTACTGGAGCTGTGGCAGACC 598
 Db 17422 TCCCAACAGGCGGCGGAGATCCAGTTGTGGCAGTTTCTACTGGAGCTGTGGCAGACC 17481

QY 599 GCGGAAAGCGCGGCTGCATCCGCTGGAGGGCGGCGGAGTTCAAGCTACCGACC 658
 Db 17482 GCGGAAAGCGCGGCTGCATCCGCTGGAGGGCGGCGGAGTTCAAGCTACCGACC 17541

QY 659 CCGACGAGGTGGCGGCGAGCTGGGCGGAGCCAGAGCAAGCCCAATATGAATACGACA 718
 Db 17542 CCGACGAGGTGGCGGCGAGCTGGGCGGAGCCAGAGCAAGCCCAATATGAATACGACA 17601

QY 719 AGCTAAGTCGAGCACTGCGCTACTACTACGACAAAAACATCATGAGCAAGGTGCACGCA 778
 Db 17602 AGCTAAGTCGAGCACTGCGCTACTACTACGACAAAAACATCATGAGCAAGGTGCACGCA 17661

QY 779 AGCGCTAGCCTACCGCTTTGACTTCCAGGCGCTGGGACAGGCTTGCAGCCACCCCG 838
 Db 17662 AGCGCTAGCCTACCGCTTTGACTTCCAGGCGCTGGGACAGGCTTGCAGCCACCCCG 17721

QY 839 CGCAGCCCGCGCGCGCTGCGCGCGCGCGGCGGAGCGCGCGCGCGGAGTGGCGCAC 898
 Db 17722 CGCAGCCCGCGCGCGCTGCGCGCGCGCGGCGGAGCGCGCGCGCGGAGTGGCGCAC 17781

QY 899 TTACAAGCTCCCGGCTGCTGCGCTCCACTGCGCTTCCCGGCGCTTCCAAACTCAACC 958
 Db 17782 TTACAAGCTCCCGGCTGCTGCGCTCCACTGCGCTTCCCGGCGCTTCCAAACTCAACC 17841

QY 959 TTATGGAGCGCTCGGCGCGCGCTGCGCGCGCTTCTTACTGCGCTTGGTCCCAAG 1018
 Db 17842 TTATGGAGCGCTCGGCGCGCGCTGCGCGCGCTTCTTACTGCGCTTGGTCCCAAG 17901

QY 1019 CCACGCGCGCTGCGCGCGCGCGCGCGCGCTGCGCTTACCACCGCGCTTGGAGCGCCCTC 1078
 Db 17902 CCACGCGCGCTGCGCGCGCGCGCGCGCGCTTACCACCGCGCTTGGAGCGCCCTC 17961

QY 1079 CCGGCGCGCTTGGCGCGCTGCGCGCGCTTCCGCACTTGGGCGGTCAATTATCACTAGACG 1138
 Db 17962 CCGGCGCGCTTGGCGCGCTGCGCGCGCTTCCGCACTTGGGCGGTCAATTATCACTAGACG 18021

QY 1139 GACGCGCGGTGACAGTGGGCGCTTCCACACAGCAGTGAACCAATCCCATCTCATCT 1198
 Db 18022 GACGCGCGGTGACAGTGGGCGCTTCCACACAGCAGTGAACCAATCCCATCTCATCT 18081

QY 1199 GGGAGGAGCGCGGAGATTTCCCGCGCGCTTACACAGATTTCCGTTGCAGCAGCGG 1258
 Db 18082 GGGAGGAGCGCGGAGATTTCCCGCGCGCTTACACAGATTTCCGTTGCAGCAGCGG 18141

QY 1259 CTCCACGCGCGGAGGAGATGGGAAGCTCTGAGGTCTTCTTCAATACAGGCTT 1318
 Db 18142 CTCCACGCGCGGAGGAGATGGGAAGCTCTGAGGTCTTCTTCAATACAGGCTT 18201

QY 1319 CCAGGCTCCCATATATCATACCCCGAGGAGGTGCATGTGCTCCCACTTTAATTTTCTC 1378
 Db 18202 CCAGGCTCCCATATATCATACCCCGAGGAGGTGCATGTGCTCCCACTTTAATTTTCTC 18261

QY 1379 TTCCAAGTCTCCAGATTTCTGGAATCCCGCTTTTCTTCTTCTTCTTCTTCTTCTTCT 1438
 Db 18262 TTCCAAGTCTCCAGATTTCTGGAATCCCGCTTTTCTTCTTCTTCTTCTTCTTCTTCT 18321

QY 1439 GCCTTCTCTTTATGACCCCTAGTTTCTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGT 1498
 Db 18322 GCCTTCTCTTTATGACCCCTAGTTTCTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGT 18381

QY 1499 TCATTTTTTCTCTCCACGACCTACTCCAAACGGTAGTACCTCGGTAGTACCTCGAGG 1558
 Db 18381 TCATTTTTTCTCTCTCCACGACCTACTCCAAACGGTAGTACCTCGGTAGTACCTCGAGG 1558

```

Db      18382 TCATTTCCTCCAGCACTACTCAAAACGGTAGTACTCGTAGTACCTCGAGG 18441
OY      1559 CTTCTCACACTCCCTTTTGGGATATGAGAAAGCATCAAAAACATCTCTGCTGTTGTCCA 1618
Db      18442 CTTCTCACACTCCCTTTTGGGATATGAGAAAGCATCAAAAACATCTCTGCTGTTGTCCA 18501
OY      1619 TCCTATCCCAACACTCTCGCTTCGGCTCCCTTCATACACACTCTGGCCCAAGGACCTT 1678
Db      18502 TCCTATCCCAACACTCTCGCTTCGGCTCCCTTCATACACACTCTGGCCCAAGGACCTT 18561
OY      1679 CGTCTGTATATATTCCTTTTCAGCCCCCATTAAGATCCAAAGCTTC 1722
Db      18562 CGTCTGTATATATTCCTTTTCAGCCCCCATTAAGATCCAAAGCTTC 18605

RESULT 4
AC095777
LOCUS   Rattus norvegicus clone CH230-9K24, *** SEQUENCING IN PROGRESS ***,
DEFINITION
3 unordered pieces.
AC095777
VERSION AC095777.6 GI:22855924
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_ENRICHED.
SOURCE  Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus
          Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
          Rattus.
1 (bases 1 to 287955)
Muzny,D.Marie., Metzker,M.Lee., Abramson,S., Adams,C., Alder,J.,
Allen,C., Allen,H., Alsbrooks,S., Amin,A., Anguiano,D.,
Anyalebechi,V., Aoyagi,A., Ayodeji,M., Baca,E., Baden,H.,
Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F.,
Biswal,D., Blair,J., Blankenburg,K., Blyth,P., Brown,M.,
Bryant,N., Buhay,C., Burch,P., Burrell,K., Calderon,E.,
Cardenas,V., Carter,K., Cavazos,I., Ceasar,H., Center,A.,
Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J.,
Cleaveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L.,
Davila,M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D.,
Delgado,O., Denson,S., Deramo,C., Ding,Y., Dinh,H., Divya,K.,
Draper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Eaves,K.,
Egan,A., Escotto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G.,
Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P.,
Fraser,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Garza,M.,
Gebregorgis,E., Geer,K., Gill,R., Grady,M., Guerrero,M., Guevara,W.,
Gunnaratne,P., Haaland,W., Hamill,C., Hamilton,C., Hamilton,K.,
Harvey,Y., Havlak,P., Hawes,A., Henderson,N., Hernandez,J.,
Hernandez,R., Hines,S., Hladun,S.L., Hodgson,A., Hogues,M.,
Hollins,B., Howells,S., Hulyk,S., Hume,J., Idlebird,D., Jackson,A.,
Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jollivet,A.,
Karpathy,S., Kelly,S., Kelly,S., Khan,Z., King,L., Kovar,C.,
Kowis,C., Kraft,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J.,
Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J.,
Lorensheewa,L., Loulseghe,H., Lozado,R.J., Lu,X., Ma,J.,
Maheshwari,M., Mahindartne,M., Mahmoud,M., Malloy,K., Mangum,A.,
Mangum,B., Mapua,P., Martin,K., Martin,R., Martinec,E.,
Mawhiney,S., McLeod,M.P., McNeill,T.Z., Meenen,E.,
Milosavljevic,A., Miner,G., Minja,E., Montanayor,J., Moore,S.,
Morgan,M., Morris,K., Morris,S., Munidasa,M., Murphy,M., Nair,L.,
Nankervis,C., Neal,D., Newton,N., Nguyen,N., Norris,S.,
Nwakoelameh,O., Okwono,G., Olarnpunsagoon,A., Pal,S., Parks,K.,
Pasternak,S., Paul,H., Perez,A., Perez,L., Pfankoch,C.,
Plopper,F., Poindexter,A., Popovic,D., Primus,E., Pu,L.-L.,
Puazo,M., Quiroz,J., Rachlin,E., Reeves,K., Regier,M.A., Reigh,R.,
Reilly,B., Reilly,M., Ren,Y., Reuter,M., Richards,S., Riggs,F.,
Rivers,C., Rodkey,T., Rojas,A., Rose,M., Rose,R., Ruiz,S.J.,
Sanders,W., Savery,G., Scherer,S., Scott,G., Shatsman,S., Shen,H.,
Shetty,J., Shvartsbeyn,A., Sisson,I., Sitter,C.D., Smajs,D.,
Sneed,A., Sodergren,E., Song,X.-Z., Sorelle,R., Sosa,J.,
Steimle,M., Strong,R., Sutton,A., Svatek,A., Tabor,P., Taylor,C.,
Taylor,T., Thomas,N., Thomas,S., Tingey,A., Trejos,Z., Usmani,K.,
Valas,R., Vera,V., Villasana,D., Waldron,L., Walker,B., Wang,J.,
Wang,Q., Wang,S., Warren,J., Warren,R., Wei,X., White,F.,

```

```

Williams,G., Willson,R., Wleczyk,R., Wooden,H., Worley,K.,
Wright,D., Wright,R., Wu,J., Yakub,S., Yen,J., Yoon,L., Yoon,V.,
Yu,F., Zhang,J., Zhou,J., Zhou,X., Zhao,S., Dunn,D., von
Niederhausern,A., Weiss,R., Smith,D.R., Holt,R.A., Smith,H.O.,
Weinstock,G. and Gibbs,R.A.
Direct Submission
Unpublished
2 (bases 1 to 287955)
Worley,K.C.
Direct Submission
Submitted (17-SEP-2001) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 287955)
Rat Genome Sequencing Consortium.
Direct Submission
Submitted (03-OCT-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On Sep 14, 2002 this sequence version replaced gi:22758717.
The sequence in this assembly is a combination of BAC based reads
and whole genome shotgun sequencing reads assembled using Atlas
(http://www.hgsc.bcm.tmc.edu/projects/rat/). As a result, the
sequence may extend beyond the ends of the clone and there may be
contigs that consist entirely of whole genome shotgun sequence
reads. Both end sequences and whole genome shotgun sequence only
contigs will be indicated in the feature table.
----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: GDBR
Center clone name: CH230-9K24
----- Summary Statistics
Assembly program: Phrap; version 0.990329
Consensus quality: 192943 bases at least Q40
Consensus quality: 197649 bases at least Q30
Consensus quality: 200917 bases at least Q20
Estimated insert size: 225772; sum-of-contigs estimation
Quality coverage: 4x in Q20 bases; sum-of-contigs estimation
-----
* NOTE: Estimated insert size may differ from sequence length
(see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 3 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
* 1 285136: contig of 285136 bp in length
* 285137 285236: gap of unknown length
* 285237 286548: contig of 1412 bp in length
* 286649 286748: gap of unknown length
* 286749 287955: contig of 1207 bp in length.
*
* Location/Qualifiers
* 1..287955
* /organism="Rattus norvegicus"
* /mol_type="genomic DNA"
* /db_xref="taxon:10116"
* /clone="CH230-9K24"
* 1..1275
* /note="wgs contig"
* misc_feature 27580..29551
* misc_feature /note="wgs_contig"
* BASE COUNT 51622 a 49444 c 48887 g 53788 t 84214 others
* ORIGIN
*
Query Match 67.2%; Score 1177.6; DB 2; Length 287955;
Best Local Similarity 99.7%; Pred. No. 5.1e-227;

```

		Matches 1180;	Conservative 0;	Mismatches 4;	Indels 0;	Gaps 0;
QY	539	TACACAAAGGAGCGGCGAGATCCAGTTGTGTCAGTTCTTACTGAGCTGCTGGCAGACC	598			
DB	150640	TCCCAACAGGACGCGGAGATCCAGTTGTGTCAGTTCTTACTGAGCTGCTGGCAGACC	150699			
QY	599	GCGCAACCGCGGCTGCATCGCTGGGAGGCGGCACGCGGAGTTCAAGCTCACCGACC	658			
DB	150700	GCGCAACCGCGGCTGCATCGCTGGGAGGCGGCACGCGGAGTTCAAGCTCACCGACC	150759			
QY	659	CGGACAGGTGGCGAGCTGGGCGAGCGGAGAGGAGCCCAATATGAACACTAGACA	718			
DB	150760	CGGACAGGTGGCGAGCTGGGCGAGCGGAGAGGAGCCCAATATGAACACTAGACA	150819			
QY	719	AGCTAAGTCGAGCACTGCGCTACTACTACGACAAAACATCATGAGCAAGGTGACGGCA	778			
DB	150820	AGCTAAGTCGAGCACTGCGCTACTACTACGACAAAACATCATGAGCAAGGTGACGGCA	150879			
QY	779	AGCGTACGCTACCGCTTTGACTTCCAGGGCTGGCACAGGCTTGCAGCACCAACCG	838			
DB	150880	AGCGTACGCTACCGCTTTGACTTCCAGGGCTGGCACAGGCTTGCAGCACCAACCG	150939			
QY	839	CGCAGCCCAACCGCGCGCTGCACCTGCCCTTCCCGGCGCTTCCAACTCAACC	898			
DB	150940	CGCAGCCCAACCGCGCGCTGCACCTGCCCTTCCCGGCGCTTCCAACTCAACC	150999			
QY	899	TTTACAAGCTCCCGGCTGGCTCCACTGCCCTTCCCGGCGCTTCCAACTCAACC	958			
DB	151000	TTTACAAGCTCCCGGCTGGCTCCACTGCCCTTCCCGGCGCTTCCAACTCAACC	151059			
QY	959	TTATGGCAGCTCGCGCGGCTGGCGCGCTGGCTTCTTACTGGCTGTGTCACACG	1018			
DB	151060	TTATGGCAGCTCGCGCGGCTGGCGCGCTGGCTTCTTACTGGCTGTGTCACACG	151119			
QY	1019	CCACGCGCTGCGCGCGCAACGCTGGCTTACCGAACCGCGCTGACGCCCTC	1078			
DB	151120	CCACGCGCTGCGCGCGCAACGCTGGCTTACCGAACCGCGCTGACGCCCTC	151179			
QY	1079	CCGGGCCCTTTGGCGGCTGGCGCGCTTGGCACTTGGGGGTCATTATCACTAGACGG	1138			
DB	151180	CCGGGCCCTTTGGCGGCTGGCGCGCTTGGCACTTGGGGGTCATTATCACTAGACGG	151239			
QY	1139	GACGCGCGGTGACGTGGGCTTCCACACAGCAGTGCACCAATCCATCTCATCT	1198			
DB	151240	GACGCGCGGTGACGTGGGCTTCCACACAGCAGTGCACCAATCCATCTCATCT	151299			
QY	1199	GGGAGGACCCGAGAGTTTCCCGACGTTCTTTTACCACAGATTTCTTCCAGCAGCG	1258			
DB	151300	GGGAGGACCCGAGAGTTTCCCGACGTTCTTTTACCACAGATTTCTTCCAGCAGCG	151359			
QY	1259	CTCCAGCCAGGGAAGGAAGTGGGAAGCTCTGAGGTCTTCTTGAATACGAGGCTT	1318			
DB	151360	CTCCAGCCAGGGAAGGAAGTGGGAAGCTCTGAGGTCTTCTTGAATACGAGGCTT	151419			
QY	1319	CCAGCTCCCATATCATCACCCAGGAAGGTGCATGTCTCCACATTTATTTTCTC	1378			
DB	151420	CCAGCTCCCATATCATCACCCAGGAAGGTGCATGTCTCCACATTTATTTTCTC	151479			
QY	1379	TTCCAAGTCTCCAGATTCTGGAATCTCCCGTCTTTTCTTCTTCTACCTGGAGCCCT	1438			
DB	151480	TTCCAAGTCTCCAGATTCTGGAATCTCCCGTCTTTTCTTCTTCTACCTGGAGCCCT	151539			
QY	1439	GCCTTCTCTTTATGACCCCTAGTTTCTGTGTTTTTTTTTTTTTTTCTCTCTCTCC	1498			
DB	151540	GCCTTCTCTTTATGACCCCTAGTTTCTGTGTTTTTTTTTTTTTTTCTCTCTCTCC	151599			
QY	1499	TCATTTTTTTTCTCTCCACAGCTACTCCAAACGGTAGTACCTCGGTAGTACCTCGAGG	1558			
DB	151600	TCATTTTTTTTCTCTCCACAGCTACTCCAAACGGTAGTACCTCGGTAGTACCTCGAGG	151659			
QY	1559	CTTCTCACACTCCCTTTTCCGGATATGAGAGGATCAAAACATCTCTGCTGTGTCCA	1618			
DB	151660	CTTCTCACACTCCCTTTTCCGGATATGAGAGGATCAAAACATCTCTGCTGTGTCCA	151719			

QY	1619	TCCCTATCCCAACACTCTGGCTTGCCTCCCTTCCATACACACTCTGGCCCAAGACCT	1678			
DB	151720	TCCCTATCCCAACACTCTGGCTTGCCTCCCTTCCATACACACTCTGGCCCAAGACCT	151779			
QY	1679	CGTCTGTATATATTCCTTTTCAGCCCCCATTAAGATCCAAAGCTTC	1722			
DB	151780	CGTCTGTATATATTCCTTTTCAGCCCCCATTAAGATCCAAAGCTTC	151823			
RESULT 5						
LOCUS	AY049085	Mus musculus ETS-domain transcription factor mRNA	1489 bp	linear	ROD 12-MAR-2002	
DEFINITION	AY049085	Mus musculus ETS-domain transcription factor mRNA, complete cds.				
ACCESSION	AY049085					
VERSION	AY049085.1	GI:19386466				
KEYWORDS						
SOURCE		Mus musculus (house mouse)				
ORGANISM		Mus musculus				
REFERENCE		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 1489)				
AUTHORS		Pfaar,H., von Holst,A., Vogt Weisenhorn,D.M., Brodski,C., Guimera,J. and Wurst,W.				
TITLE		mPet-1, a mouse ETS-domain transcription factor, is expressed in central serotonergic neurons				
JOURNAL		Dev. Genes Evol. 212 (1), 43-46 (2002)				
MEDLINE		21863846				
PUBMED		11875656				
REFERENCE		2 (bases 1 to 1489)				
AUTHORS		Pfaar,H., von Holst,A., Vogt Weisenhorn,D.M., Brodski,C., Guimera,J. and Wurst,W.				
TITLE		Direct Submission				
JOURNAL		Submitted (31-JUL-2001) Molecular Neurogenetics, Max-Planck-Institute of Psychiatry, Kraepelinstrasse 2-10, Munich, Bavaria 80804, Germany				
FEATURES		Location/Qualifiers				
source		1. 1489				
		/organism="Mus musculus"				
		/mol_type="mRNA"				
		/strain="C3H"				
		/db_xref="taxon:10090"				
		209. .922				
		/note="mPet-1"				
		/codon_start=1				
		/product="ETS-domain transcription factor"				
		/protein_id="AAL13055.1"				
		/db_xref="GI:19386467"				
		/translation="MROSGTSQPLLINMYLPDPVGDGLFKBKSWSGSLSPAVOKGS				
		GOIOLWQFLLELLADRANAGCIAWEGHGEFLTPDEVARRWGRKSKPNWYDKLS				
		RALRYVDKNTMSVHGKRYAYRFQGLAQACQPPPAHAHAHAHAHAHAHAHAHAHA				
		YKLPAGLAPLPFPGLSKLNLMASAGVAPAGFSYWPGNPATAAATAALYPTPLGP				
		PPGPFPAVAASHLGGHYH"				
BASE COUNT		262 a 562 c 378 g 287 t				
ORIGIN						
Query Match		64.4%; Score 1129; DB 10; Length 1489;				
Best Local Similarity		89.5%; Pred. No. 2.8e-217;				
Matches 1338; Conservative 0; Mismatches 120; Indels 37; Gaps 10;						
QY	231	GCTCCAGACCCCTCTCCCTCCACACACCTGGCTCCCGCTCCCGCAGGCTTGGC	290			
DB	18	GCTCCAGACCCCTCTCCCTCCCGCCACACCTGGCTCCCGCTCTCCAGGCTTGGC	77			
QY	291	GGACCCGCGGCTGTACTTCTTCTCCCTGTCAACCCACAGTACACGCGGGGTATCGGCAC	350			
DB	78	GGACCCGCGGCTGTACTTCTTCTCCCTGTCAACCCACAGTACACGCGGGGTATCGGCAC	137			
QY	351	CCCAAGCGCAAGCTGACGTCG-CCCCCGCTGGTCCCCCCCATCTCCACCGCCGAGT	409			
DB	138	CCCAAGCGCAAGCTGACGTCG-CCCCCGCTGGTCCCCCCCATCTCCACCGCCGAGT	197			
QY	410	CCCCCGGCGGATGAGACAGACGCGGACCTCCCGACCCCTGCTGATCAACATGTACTAC	469			

Db	198	CCCCGGCAGCGATGAGACAGAGCGGCACCTCCCGAGCCCTGCTGATCAACATGACTCTGC	257
Qy	470	CAGATCCCGTCGGAGATGGTCTTTTAAAGAAAGGAAAGACCCGAGCTGGCGCCGCTGA	529
Db	258	CAGATCCCGTCGGAGATGGTCTTTTAAAGAAAGGAAAGACCCGAGCTGGCGCCGCTGA	317
Qy	530	GCCTCGGTACAGAAAGGAGCGGGCAGATCCAGTTGTGGCAGTTTCTACTGGAGCTGC	589
Db	318	GCCTCGGTACAGAAAGGAGCGGGCAGATCCAGTTGTGGCAGTTTCTACTGGAGCTGC	377
Qy	590	TGCGAGACCGCGGAACCGCGCTGCATCCGTTGGAGGGCGCCAGCGGAGCTTCAAGC	649
Db	378	TGCGAGATCCGCGGAACCGCGCTGCATCCGTTGGAGGGCGCCAGCGGAGTTTCAAGC	437
Qy	650	TCACCGACCCGACGAGGTGGCGACGCTGGGGCGAGCGCAAGAGCAAGCCCAATATGA	709
Db	438	TCACCGACCCGACGAGGTGGCGCGCTGGGGCGAGCGCAAGAGCAAGCCCAATATGA	497
Qy	710	ACTACGACAAGCTAAGTCGAGCACTGGCTACTACTACGACAAAGACATCATGAGCAAGG	769
Db	498	ACTACGACAAGCTAAGTCGAGCACTGGCTACTACTACGACAAAGACATCATGAGCAAGG	557
Qy	770	TGCAGCGCAAGCGCTACGCTTACGCTTTTACGTTTCCAGGGCTGGCACAGGCTTGCACG	829
Db	558	TGCAGCGCAAGCGCTACGCTTACGCTTTTACGTTTCCAGGGCTGGCACAGGCTTGCACG	617
Qy	830	CACCACCCGCGCACGCCACCGCGCTGCCCGCGCGAGCGGACGCGCGCCGACG	889
Db	618	CACCACCCGCGCACGCCACCGCGCTGCCCGCGCGAGCGGACGCGCGCCGACG	677
Qy	890	ATGGCGCACTTACAGCTCCGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGG	949
Db	678	ACGGCGCCCTCTACAAAGCTCCCGCGCGGCTGGCGCGCTGGCTGGCTGGCTGGCTGGCTGG	737
Qy	950	AACTCAACCTTATGGCAGCGCTCCGCGCGCTGGCGCGCTGGCTGGCTGGCTGGCTGGCTGG	1009
Db	738	AACTCAACCTTATGGCAGCGCTCCGCGCGCTGGCGCGCTGGCTGGCTGGCTGGCTGGCTGG	797
Qy	1010	GTCCCAACGCAACCGCGCTGCGCGCGCAACCGCTGGCTGGCTGGCTGGCTGGCTGGCTGG	1069
Db	798	GCCCGAAGCGCACCGCGCTGCCCGCGCACCGCGCGCTCTACCCCGCCCGCGCTTGC	857
Qy	1070	AGCCCTCCCGCGCGCTTTGGCGCGTGGCGCGCTGGCGCGCTGGCGCGCTGGCGCGCTGGCG	1129
Db	858	AGCCCTCCCGCGCGCTTTGGCGCGTGGCGCGCTGGCGCGCTGGCGCGCTGGCGCGCTGGCG	917
Qy	1130	ACTAGAC-GGGACGGCGGGTGAGTGGGGCTCTCCACACAG-CCAGTGACCAATCCC	1187
Db	918	ACTAGACGGGACGGCGGGTACAGCGGGCTCTCTCACATGGCCAGTGACCAATCCC	977
Qy	1188	ATCCTCATCTCGGAGAGCGCCGAAAGATTCCCGCGAGTTCCCTTTACCAACAGATTTCGT	1247
Db	978	ATCCTCATCTCGGAGAGCGCCGAAAGATTTCCTCCGACGTTCTTTAAACACAGATTTCAT	1037
Qy	1248	TGCAGACGCGCTCCAGCCACGGAGAAAGATGGAGCGCTCTGAGGTCTTCCTTGA	1307
Db	1038	TGCAGACGCGCTCCAGCCACGGAGAAAGATGGAGAAACCCCGAGGTCTTCCTTGA	1097
Qy	1308	ATACGAGGCTTCAGGCTCCCATTTATCATCACCCCGAGAGGGTGCATGTGCTCCCACTT	1367
Db	1098	ACATCAGGCTTCAGGC-CCCGTTATCTTCGCGCCCGAGGAGGTGCATGGCTCCCACTT	1156
Qy	1368	TAATTTTCTCTCCAGTCTCAGATCTCGAATCTCGAATCTCCCGTCTTTTTTCTCTCTCAC	1427
Db	1157	TAATTTTCTCTCTCCAGTCTCAGATCTCGAATCTCGAATCTCCCGT-CTCTTTTCTTTCTCAC	1215
Qy	1428	CTGAGCGCTGCTCTCTTTATGACCCCTAGTTTCTGTTTCTGTTTCTTTTCTTTTCTTCC	1487
Db	1216	TTGCACCCC---TGCACCTCGTTATGACCCCTATTTTCTGGGATTTCTTGTGCTCTCAT	1272
Qy	1488	TTCTCTCTCTCTCAATTTTTTTCTCTCCACGACCTACTCCAAACGCTAGTACTCTCGGTA	1547
Db	1273	TTTCTGAC-----TCTCTCCCAACCCCACTCCAAA-----AGTA	1308
Qy	1548	GTACCTCGAGGCTTTCACACTCCCTTT-----TCGGGATATGAGAGCATC-AAAAACA	1602
Db	1309	GTACCTCAAGGCTTTCGCACTCCCTATTTCTCTCGGATATGAGAGCTTCGATAACA	1368
Qy	1603	TCTCTGCTGTTCATCCCTATCCCAACTCTGCTTCGCTCCCTTCATACCACT	1662
Db	1369	TCTCTGCTGTTCATCCCTATCCCAACTCTGCTTCGCTCCCTTCATACCACT	1428
Qy	1663	CTGGCCCAAGGACCTCGCTGTATATATCTCTTTCAGCCCATTAAGATCCAA	1717
Db	1429	CTGGCCCAAGGACCTCGCTGTATATATCTCTTTCAGCCCATTAAGATCCCA	1483
RESULT 6			
AY049086			
LOCUS	AY049086	4956 bp	DNA linear ROD 12-MAR-2002
DEFINITION	Mus musculus ETS-domain transcription factor gene, complete cds.		
ACCESSION	AY049086		
VERSION	AY049086.1	GI:19386464	
KEYWORDS			
SOURCE	Mus musculus (house mouse)		
ORGANISM	Mus musculus		
REFERENCE	1 (bases 1 to 4956)		
AUTHORS	Pfaar,H., von Holst,A., Vogt Weisenhorn,D.M., Brodski,C.,		
TITLE	mPet-1, a mouse ETS-domain transcription factor, is expressed in		
JOURNAL	Central serotonergic neurons		
MEDLINE	Dev. Genes Evol. 212 (1), 43-46 (2002)		
PUBMED	11875656		
REFERENCE	2 (bases 1 to 4956)		
AUTHORS	Pfaar,H., von Holst,A., Vogt Weisenhorn,D.M., Brodski,C.,		
TITLE	Guimera,J. and Wurst,W.		
JOURNAL	Direct Submission		
REFERENCE	Submitted (31-JUL-2001) Molecular Neurogenetics,		
JOURNAL	Max-Planck-Institute of Psychiatry, Kraepelinstrasse 2-10, Munich,		
FEATURES	Bavaria 80804, Germany		
source	1. .4956		
	/organism="Mus musculus"		
	/mol_type="Genomic DNA"		
	/strain="129/SvJ"		
	/db_xref="taxon:10090"		
mRNA	join(1041..1302,1980..2054,3803..4956)		
CDS	/product="ETS-domain transcription factor"		
	join(1251..1302,1980..2054,3803..4389)		
	/notes="mPet-1"		
	/codon_start=1		
	/product="ETS-domain transcription factor"		
	/protein_id="AAL13056.1"		
	/db_xref="GI:19386465"		
	/translation="MRQSGTSPLLINMYLPDVGDLFKGKSPWGLSPAVQKGS		
	GGIQWQFLELLADRANAGCIAGWGHGFKLTDPEVARWRRGSRKSPNMYDKLS		
	RALRYYYDKNINSKVHGRYAYRFQGLAQCPPPAHAAHAAAAAQAQDGL		
	YKLPAGLAPLPFGGLSKLNMAASAGVAPGFSWPGPNATAAATAALYPTPLQP		
	PPGPGFAVAASHLGHHYH"		
BASE COUNT	1037 a	1418 c	1432 g 1065 t
ORIGIN	4 others		
Query Match	48.0%; Score 841.4; DB 10; Length 4956;		
Best Local Similarity	87.2%; Pred. No. 2.9e-159;		
Matches 1037; Conservative	0; Mismatches 116; Indels 36; Gaps 9;		
Qy	536	CGGTACAGAAAGGACGCGGCGAGATCCAGTTGTGGCAGTTTCTACTGGAGCTCTGGCAG	595
Db	3791	CGTTCCCGACAGCGCGGCGAGATCCAGTTGTGGCAGTTTCTCTCGGAGCTCTGGCAG	3850
Qy	596	ACGCGCGCAACCCGCTCGCATCGTGGGAGGGCGCCAGCGGAGTTCAAGCTCACCG	655

----- Project Information

Center project name: L28974
 Center clone name: 405_L_19
 ----- Summary Statistics
 Sequencing vector: Plasmid; n/a; 100% of reads
 Chemistry: Dye-terminator Big Dye; 100% of reads
 Assembly program: Phrap; version 0.960731
 Consensus quality: 186050 bases at least Q40
 Consensus quality: 187399 bases at least Q30
 Consensus quality: 187949 bases at least Q20
 Insert size: 175000; agarose-fp
 Insert size: 188187; sum-of-contigs
 Quality coverage: 8.8 in Q20 bases; agarose-fp
 Quality coverage: 8.2 in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently consists of 9 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

* 50283: contig of 50283 bp in length
 * 50284: gap of 100 bp
 * 50384: contig of 1105 bp in length
 * 51488: gap of 100 bp
 * 51489: gap of 100 bp
 * 51589: contig of 3426 bp in length
 * 55015: gap of 100 bp
 * 55115: contig of 2641 bp in length
 * 57755: gap of 100 bp
 * 57855: contig of 7209 bp in length
 * 65064: gap of 100 bp
 * 65065: gap of 100 bp
 * 65165: gap of 5111 bp in length
 * 70475: gap of 100 bp
 * 70575: gap of 100 bp
 * 158522: contig of 87947 bp in length
 * 158523: gap of 100 bp
 * 179423: contig of 20805 bp in length
 * 179428: gap of 100 bp
 * 179528: 188987: contig of 9460 bp in length.

FEATURES

source

1. .188987
 /organism="Mus musculus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10090"
 /chromosome="1"
 /map="1"
 /clone="RP24-405L19"
 /clone_lib="RPCI-24 Male Mouse BAC"

misc_feature

1. .50283
 /notes="assembly_fragment"
 clone end:SP6
 vector_side:left"

misc_feature

50384. .51488
 /notes="assembly_fragment"

misc_feature

51589. .55014
 /notes="assembly_fragment"

misc_feature

55115. .57755
 /note="assembly_fragment"

misc_feature

57856. .65064
 /notes="assembly_fragment"

misc_feature

65165. .70475
 /notes="assembly_fragment"

misc_feature

70576. .158522
 /notes="assembly_fragment"

misc_feature

158623. .179427
 /notes="assembly_fragment"

misc_feature

179528. .188987
 /notes="assembly_fragment"

misc_feature

clone end:T7
 vector_side:right"

BASE COUNT

50833 a 44352 c 45081 g 47920 t 801 others

ORIGIN

Query Match 48.0%; Score 841.4; DB 2; Length 188987;
 Best Local Similarity 87.2%; Pred. No. 3.2e-159;
 Matches 1037; Conservative 0; Mismatches 116; Indels 36; Gaps 9;

Qy 536 CGGTACAGAAAGCAGCGCGGCAGATCCAGTTGTGGCAGTTTCTACTGGAGCTGCTGGCAG 595
 Db 67936 CGTTCCCCACAGCAGCGCGGCAGATCCAGTTGTGGCAGTTTCTCTGGAGCTGCTGGCAG 67995

Qy 596 ACCGCGGAACGCGCGCTGCTGCGTGGGAGGCGGCACGCGGAGTTCAAGCTCACCG 655
 Db 67996 ATCGCGGAACGCGCGTGGCATCGGTGGAGGCGGCACGCGGAGTTCAAGCTCACCG 68055

Qy 656 ACCCGACGAGGTGGCGCGCAGCTGGGGCGAGCGCAAGAGCAAGCCCAATATGAACACG 715
 Db 68056 ACCCGACGAGGTGGCGCGCAGCTGGGGCGAGCGCAAGAGCAAGCCCAATATGAACACG 68115

Qy 716 ACAAGCTAGTCAGCAGCTGCGGTACTACTACACAAAACATCATGAGCAAGTGCACG 775
 Db 68116 ACAAGCTAGTCGCGCGCTGCGTACTACTACACAAAACATCATGAGCAAGTGCACG 68175

Qy 776 GCAAGCGCTACGCTACCGCTTTGACTTCCAGGCGCTGGGCACAGGCTTGCAGCACACAC 835
 Db 68176 GCAAGCGCTACGCTACCGCTTTGACTTCCAGGCGCTGGGCAGGCTTGCAGCACACAC 68235

Qy 836 CCGCGCACGCCACGCGCGCGCTGCGCGCCGCGAGCGGCGCGCGCCAGGATGGCG 895
 Db 68236 CCGCGCACGCCACGCGCGCGCTGCGCGCCGCGCGCGCGCGCGCCAGGATGGCG 68295

Qy 896 CACTTTACAGCTCCCGCGCTGCTGGCTCCACTGCGCTTCCCGGCGCTCTCCAACTCA 955
 Db 68296 CCTCTACAGCTCCCGCGCGCGCTGCGCGCGCTTCCCGGCGCTTCCAACTCA 68355

Qy 956 ACCTTATGCGAGCTCGCGCGCGCTGCGCGCGCTGCTGCTTCTACTGCGCTGGTCCCA 1015
 Db 68356 ACCTTATGCGAGCTTGGCGCGCGCTTGGCGCGCGCTGCTGCTTCTACTGCGCTGG 68415

Qy 1016 ACGCCACCGCGCTGCGCGCGCGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1075
 Db 68416 ACGCCACCGCGCTGCGCGCGCGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTG 68475

Qy 1076 CTCCCGCGCGCTTGGCGCGCGCTGCGCGCGCTTGGCGCGCTTGGCGCGCTTGGCGCG 1135
 Db 68476 CTCCCGCGCGCTTGGCGCGCGCTTGGCGCGCGCTTGGCGCGCTTGGCGCGCTTGG 68535

Qy 1136 C-GGGACGCGCGGCTGAGTGGGCGCTCTCCACACAG -CCAGTGACCAATCCCATCTC 1193
 Db 68536 CGGGGACGCGCGGTTACAGCGGGGCTCTCTCACATGGCCCGAGTACCAATCCCATCTC 68595

Qy 1194 ATCTGGGAGGACCGCGAGATTTCCCGAGCGTTCTCTTTACACAGATTTCTGTTGAGC 1253
 Db 68596 ATCTCTGGAGGACCGCGAGATTTCTCCGAGCGTTCTCTTTTAAACAGATTTCTG 68655

Qy 1254 AGCCGCTCCAGCGCGCGGAGGAGGATGGGAGGCTCTGAGGCTTCTCTTGAATACGA 1313
 Db 68656 AGCGCGCGCGCGCGCGGAGGAGGATGGGAGGCTCTCTTGAATACGA 68715

Qy 1314 GGCTTCCAGGCTCCCATTTATCATCAACCCAGGAGGCTGATGCTGCCACTTTAATTT 1373
 Db 68716 GGCTTCCAGGCTCCCATTTATCATCAACCCAGGAGGCTGATGCTGCCACTTTAATTT 68774

Qy 1374 TTCTCTTCAAGTCTCCAGATTTCTGGAATCTCGGCTTTTCTTCTCTTCTACCTGGAG 1433
 Db 68775 TTCTCTTCAAGTCTCCAGATTTCTGGAATCTCGGCTTTTCTTCTCTTCTACCTGGAG 68833

Qy 1434 CCCCTGCGCTCTCTTTATGACCCCTAGTTTCTGTTTGTGTTTCTTTTCTCTCTCTC 1493
 Db 68834 CCC---TGCACTCTGTTATGACCCCTATTTTCTGGGATTTCTTGGTCTCTCTCT 68890

Qy 1494 TCTCTCTATTTTTTCTCTCCACGACCTACTCCAAACGGTAGTACCTGGGTAGTACCT 1553
 Db 68891 AC-----TCTCTCCCAACCACTCCAAA-----AGGTAGTACCT 68926

Qy	1554	CGAGGCTTCTCACACTCCCTCCCTTT-----TCGGGATATGAGAAGCATC-AAAAAACATCTCTG	1608
Db	68927	CAAGGCTTCTCGCAGCATCCCTTATTTCTCCTCGGATATGAGAAGCTCGATAAAACATCTCTG	68986
Qy	1609	CTGTGTGTCATCCCTATCCCAACACTCTGGCTTCGCTCCCTTCACATACACACTCTGGCC	1668
Db	68987	CGGTTCCTCCCAATCCCTATATCAACACTCTGGCTTCCCTTTCTCTCCCATATCACACTCTGGCC	69046
Qy	1669	CAAGGACCCCTCGTCTCTATATATTCCTTTTCAGCCCCCATTAAGATCCAA	1717
Db	69047	CAAGGACCCCGAGCTGTATATATCCCTTTTCAGCCCCCATTAAGCTCCCA	69095
RESULT 8			
AC104542			
LOCUS	AC104542		
DEFINITION	Mus musculus clone RP23-165D11, WORKING DRAFT SEQUENCE, 9 ordered pieces.		
ACCESSION	AC104542		
VERSION	AC104542.3	GI:29164653	
KEYWORDS	HTG; HTGS PHASE2; HTGS DRAFT; HTGS_FULLTOP.		
SOURCE	Mus musculus (house mouse)		
ORGANISM	Mus musculus		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
AUTHORS	1 (bases 1 to 196584)		
TITLE	Birren,B., Nusbaum,C. and Lander,E.		
JOURNAL	Mus musculus, clone Rp23-165D11		
REFERENCE	2 (bases 1 to 196584)		
AUTHORS	Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S., Barna,N., Bastien,V., Boguslavskiy,L., Boukhalter,B., Brown,A., Camarata,J., Campopiano,A., Chang,J.J., Chazaro,B., Choquel,Y., Collangelo,M., Collins,S., Collamore,A., Cook,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S., Fato,S., Ferreir,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S., Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N., Hages,B., Heatford,A., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kamat,A., Karatas,A., Kells,C., LaRocque,K., Lamazares,R., Landers,T., Lehoczy,J., Levine,R., Liu,G., MacLean,C., Macdonald,P., Major,J., Marquis,N., Matthews,C., McCarthy,M., McEwan,P., McKernan,K., McPheeters,R., Meldrim,J., Meneus,L., Mihova,T., Mienga,V., Murphy,T., Naylor,J., Nguyen,C., Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V., Raymond,C., Retta,R., Riback,M., Riley,R., Rise,C., Rogov,P., Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupack,R., Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N., Strauss,N., Subramanian,A., Talamas,J., Testaye,S., Theodore,J., Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.		
TITLE	Direct Submission		
JOURNAL	Submitted (13-DEC-2001) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA		
REFERENCE	3 (bases 1 to 196584)		
AUTHORS	Birren,B., Nusbaum,C., Lander,E., Abouelleil,A., Allen,N., Anderson,S., Arachchi,H.M., Barna,N., Bastien,V., Bloom,T., Boguslavskiy,L., Boukhalter,B., Camarata,J., Chang,J.J., Chospel,Y., Collamore,A., Cook,A., Cooke,P., Corum,B., DeArellano,K., Diaz,J.S., Dodge,S., Dooley,K., Dorris,L., Erickson,J., Fato,S., Ferreira,P., FitzGerald,M., Gage,D., Galagan,J., Gardyna,S., Graham,L., Grand-Pierre,N., Hafez,N., Hagopian,D., Hages,B., Hall,J., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kamat,A., Karatas,A., Kells,C., Landers,T., Levine,R., Lindblad-Toh,K., Liu,G., Lui,A., Mabbitt,R., Maclean,C., Macdonald,P., Major,J., Manning,J., Matthews,C., McCarthy,M., Meldrim,J., Meneus,L., Mihova,T., Mienga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Rachupka,A., Ramasamy,U., Raymond,C., Retta,R., Rise,C., Rogov,P., Roman,J., Schauer,S., Schupack,R., Seaman,S., Severy,P., Smith,C., Spencer,B., Stange-Thomann,N., Stojanovic,N., Stubbs,M.,		

misc_feature	32893..38727	/note="assembly_fragment"	
misc_feature	38828..129564	/note="assembly_fragment"	
misc_feature	129665..147635	/note="assembly_fragment"	
misc_feature	147736..184494	/note="assembly_fragment"	
misc_feature	184595..196584	/note="assembly_fragment"	
	clone_end:T7		
	vector_side:right		
BASE COUNT	51153 a 46239 c 45505 g 52883 t	804 others	
ORIGIN			
Query Match	48.0%;	Score 841.4;	DB 2; Length 196584;
Best Local Similarity	87.2%;	Pred. No. 3.2e-159;	
Matches 1037;	Conservative 0;	Mismatches 116;	Indels 36; Gaps 9;
Qy	536	CGGTACAGAAAGCAGCGGCAGCATCCAGTTGTGGCAGTTTCTACTGGAGTGTCTTCTACTGGAGTGTCTTCTCTGGAGTGTCTGGCAG	595
Db	35727	CGTTCCCCACAGCAGCGGCAGATCAGTTGTGGCAGTTTCTCTGGAGTGTCTTCTCTGGAGTGTCTGGCAG	35786
Qy	596	ACCGCGCGAAGCGCGCTGCATCGCTGGAGGGCGGCACGGCGAGTTTCAAGCTCACCG	655
Db	35787	ATCGCGGAACGCGGTTCATCCGCTGGAGGGCGGCACGGCGAGTTTCAAGCTCACCG	35846
Qy	656	ACCCGACGAGGTGGCGCGACGCTGGGGCGAGCGCAAGCAGGCCCAATATGAACCTACG	715
Db	35847	ACCCGACGAGGTGGCGCGCGCTGGGGCGAGCGCAAGCAGGCCCTAACATGAACCTACG	35906
Qy	716	ACAAGCTTAAGTCGAGCAGCTGCGCTACTCTACGACAAAAACATCATGACGAAGGTGCACG	775
Db	35907	ACAAGCTTAAGTCGCGCGCTGCGCTACTACTAGCAGAAAAACATCATGACGAAGGTGCACG	35966
Qy	776	GCAAGCGCTAGCGCTACCGCTTTGACTTCAGGGCTTGCGCAGGGCTTGCAGGCCACAC	835
Db	35967	GCAAGCGCTAGCGCTACCGCTTTGACTTCAGGGCTTGCGCAGGGCTTGCAGGCCACAC	36026
Qy	836	CCGCGCAGCGCCACGCGCGCTGCGCGCGCGCGCAGCGGCAGCGCGCCAGGAGTGGCG	895
Db	36027	CCGCGCAGCGCCACGCGCGCTGCGCGCGCGCGCGCGCTGGCGCGCCAGGAGCGCG	36086
Qy	896	CACTTTACAAGCTCCCGGTGTGTCTGGCTTCATGTCCTTCCCGCGCTTCCCAAATCA	955
Db	36087	CCCTCTACAAGCTCCCGCGCGCTGGCGCGCGCTGCGCTTCCCGCGCGCTATCCAAACTCA	36146
Qy	956	ACCTTATGGCAGCTTCGGCGCGGTGGCGCGCGCTTCTTACTGGCTGTGTCCCA	1015
Db	36147	ACTTGTAGCAGATTTCGGCGGGCTTGGCGCGCGCTTCTTCTACTTGGCTGTGGCGCA	36206
Qy	1016	ACGCCACCGCGGTGGCGCGCCACCGCTTCACCTTACCCAAACCCCGGGCTTGCAGCCCC	1075
Db	36207	ACGCCACCGCGGTGGCGCGCCACCGCGCGCTTACCCCAACCCCGGGCTTGCAGCCCC	36266
Qy	1076	CTCCGGGCGCTTTGGCGGGTGGCGCGCTTCCGACTTGGGGGTCAATTACTACTAGA	1135
Db	36267	CTCCGGGCGCTTTCGGCGCAGTGGCGGAGCTTGGCAGCTTGGGGGTCAATTACTACTAGA	36326
Qy	1136	C-GGGACGGCCGGTGCAGTGGGGCTCTCCACACAG--CCAGTGACCAATCCCATCTCTC	1193
Db	36327	CGGGGACGGCCGGGTACAGCGGGCGCTCTCTCACATGGGCCAGTGACCAATCCCATCTCTC	36386
Qy	1194	ATCCTGGGAGGCGCCGAGATTCCCGCAGGTTCCITTTACACAGATTTCGTGTGACG	1253
Db	36387	ATCCTCGGAGGAGCCCGAAGCTTCCCGCAGCTTCCCTTTAACACAGATTTTCAATGACG	36446
Qy	1254	AGCCGCTCCAGCCACAGGAAAGAAAGATGGGAAGCCTCTGAGGTCTTCTCTTGAATACGA	1313
Db	36447	AGCCGCGCCAGCCACAGGGAAGAAAGATGGGAAGATTTCTCTTGAACATCA	36506
Qy	1314	GGCTTCAGGCTCCCATTTATCATCACCCGAGGAGGTGCATGTGCTCCCACTTTAATTT	1373

Db 143644 CAAGGCTTCGCACTCCCTCTATTTCTCGGGATATGAGAGCCTCGATAAACAATCTCTG 143585
 Qy 1609 CTGTTGTCATCCCTATCCCAACACTCTGGCTTCGGCTCCCTTCATACACACTCTGGCC 1668
 Db 143584 CCGTTCCCATCCCTATATCAACTCTGGCTTCCTTCCTCCCATATCAACTCTGGCC 143525
 Qy 1669 CAAGGACCTCGTCTGTATATATTCCTTTAGCCCCCATTAAGATCCAA 1717
 Db 143524 CAAGGACCCAGCCTGTATATATCCCTTTAGCCCCCATTAAGATCCCA 143476

RESULT 11

BC023511
 LOCUS BC023511 1805 bp mRNA linear PRI 23-SEP-2002
 DEFINITION Homo sapiens, FEV protein, clone MGC:12793 IMAGE:4130242, mRNA, complete cds.
 ACCESSION BC023511
 VERSION BC023511.1 GI:23270720
 KEYWORDS MGC.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 1805)
 Strausberg, R.
 Direct Submission
 Submitted (05-FEB-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

REMARK

NIH-MGC Project URL: <http://mgc.nci.nih.gov>
 Contact: MGC help desk
 Email: cgapbs-r@mail.nih.gov
 Tissue Procurement: ATCC
 cDNA Library Preparation: Rubin Laboratory
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: National Institutes of Health Intramural Sequencing Center (NISC), Gaithersburg, Maryland;
 Web site: <http://www.nisc.nih.gov/>
 Contact: nisc.mgc@hghri.nih.gov
 Akhter, N., Ayale, K., Beckstrom-Sternberg, S.M., Benjamin, B., Blakesley, R.W., Bouffard, G.G., Breen, K., Brinkley, C., Brookes, S., Dietrich, N.L., Granite, S., Guan, X., Gupta, J., Haghghi, P., Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Latic, P., Legaspi, R., Maduro, Q.L., Masiello, C., Maskeri, B., Mastrian, S.D., McCloskey, J.C., McDowell, J., Pearson, R., Stantripop, S., Thomas, P.J., Touchman, J.W., Tsurgon, C., Vogt, J.L., Walker, M.A., Wetherby, K.D., Wiggins, L., Young, A., Zhang, L.-H. and Green, E.D.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Series: IRAL plate: 17 Row: i Column: 5
 This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 8923788.

FEATURES

source

1. 1805
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="MGC:12793 IMAGE:4130242"
 /tissue_type="Brain, neuroblastoma"
 /clone_lib="NIH MGC 19"
 /lab_host="DH10B-R"
 /note="Vector: pOTB7"
 487, 1203
 /codon_start=1
 /product="FEV protein"
 /protein_id="AAH23511.1"
 /db_xref="GI:23270721"
 /db_xref="LocusID:54738"
 /translation="MRQSGASQPLINMVLDPDVGDLFKDGNKPSWGLPSPAVQKGS
 GQIQWQFLLELLADRANAGCIAWEGHGGEFKLTPDPEVARRWGERKSKPMNYDKLS

CDS

Query Match 43.6%; Score 764.4; DB 9; Length 1805;
 Best Local Similarity 73.1%; Pred. No. 9.5e-144;
 Matches 1110; Conservative 0; Mismatches 381; Indels 27; Gaps 9;
 BASE COUNT 250 a 681 c 532 g 302 t
 ORIGIN
 Qy 240 CCCCCTCCCTCCTCACCACACACCCCTGCGCGGCTCCCGCCAGGCTTCGCGGACCCGGC 299
 Db 304 CCCCCTCAACCCCGCCACGCGCGGCTGCGCCCTCCCGCCAGGCTTCGCGGACCCGGC 363
 Qy 300 GGCGTCTACTCTTCCCTGTACCCACACAGTACACAGCGCGGTATCGGCACC-CCAAGCG 358
 Db 364 GCGGTCTTCTCTCTCTGTACCCCGGTCTCTCGGGCGGGGATCGGTGCCACGAGCG 423
 Qy 359 CAAAG-CTGACGTGCCCGCGTCCCGGCTCCCGCCCATCTCCACCGCCAGTCCCCGGCA 417
 Db 424 CAAAGCCTGCTCGCCCGCTTCCCGCTCCCGCCCATCTCCACCGCCAGTCCCCGGCG 483
 Qy 418 GCGATGACAGACGCGGACCTCCAGCCCTGCTGTGATCAACATGTACTACAGATCCC 477
 Db 484 GCGATGACAGACGCGGCTCCAGCCCTGCTGTGATCAACATGTACTACAGATCCC 543
 Qy 478 GTCCGAGATGCTTTTAAAGGAAGGAAGAGCCGAGCTCGGGGCGCTGAGCCCTCGG 537
 Db 544 GTCGAGACGCTCTCTCAAGGACGGGAAGAACCGAGCTGGGGGCGCTGAGCCCGCG 603
 Qy 538 GTACAGAAAGGACGCGGCGAGATCCAGTTGTGGCAGATTTCTTACTGGAGCTGTGCGAG 597
 Db 604 GTTCAGAAAGGACGCGGACAGATCCAGTGTGGCAGATTTCTTCTGGAGCTGTGCGTAC 663
 Qy 598 CGCGGAAAGCGCGCTGATCGCTGGAGGGCGGCGCACGGCGAGTTCAAGTCTACCGAC 657
 Db 664 CGCGGAAAGCGCGCTGATCGCTGGAGGGCGGCTCACGGCGAGTTCAAGTCTACCGAC 723
 Qy 658 CCCGACGAGGTGGCGGCGACGCTGGCGGAGCGGAAGAGCAAGCCCAATATGAACCTACGAC 717
 Db 724 CCAGACGAGGTGGCGGCGGCGGTGGGGGAGCGCAAGAGCAAGCCCAATATGAACCTACGAC 783
 Qy 718 AAGCTAAGTCGAGACATCGCGTACTACTACGACAAAAACATCATGAGCAAGGTGACCGCG 777
 Db 784 AAGCTGAGCGCGCGCTGCGCTACTACTACGACAAAGAACATCATGAGCAAGGTGATGCG 843
 Qy 778 AAGCGCTACGCTTACCGCTTTGACTTCAGGGCTGTGGCAGAGCTTGCAGGCCACACCC 837
 Db 844 AAGCGCTACGCTTACCGCTTTGACTTCAGGGCTGTGGCAGAGCTTGCAGGCCCGCGCC 903
 Qy 838 GCGCACGCCACGCGCGCGCTGCGCGCGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCA 897
 Db 904 GCGCACGCTCATGCGCGCGCGAGCTGCTGCCCGCGCGCGCGCGCGCGCGCGCGCGCG 963
 Qy 898 CTTTACAAGCTCCCGGCTGGTCTGGCTCCACTGCGCTTTCCCGGCTCTCCCAACTCAAC 957
 Db 964 CTCTACAAGCTCCCGCGCGCGCTGCGCGCTGCGCTTTCCCGGCTCTCCCAACTCAAC 1023
 Qy 958 CTTATGCGAGCTTCGCGCGCGGTGGCGCGCGCTTCTTCTTACTGGCTTGGTCC--C 1014
 Db 1024 CTCATGCGCGCTCGCGCGGGGTGCGCGCGCGCGCTTCTCTACTGCGCGCGCGCGCG 1083
 Qy 1015 AAGCCACCGCGCTGCGCGCGCGCGCTGCGCTTACCCCAACCGCGCGCTTGCAGCGCC 1074
 Db 1084 CCGCGCGCGCGCTGCGCGCGCGCGCTTACCGCGCGCTTACCGCGCGCTTGCAGCGCC 1143
 Qy 1075 CTTCCCGGCGCTTTGGCGGGTGGCGCGCGCTTCGCACCTTGGGGGGTCAATTATCACTAG 1134
 Db 1144 CCGCGCGCGCTTTCGGGGCGGTGGCGCGCGCTGCGACTTGGGGGGCCATTACCACTAG 1203
 Qy 1135 ACGGAGCGCGGTGAGTGGGGCTTCTCCACACAGCCAG-----TGACCAATCCCAT 1189
 Db 1204 ACGGGCGGTGCGGTGC-CTGCGCGCTCGCGCGCGCGCTAGAGTCTCGCGCGCATCCAT 1262

JOURNAL REFERENCE AUTHORS TITLE JOURNAL	JOURNAL REFERENCE AUTHORS TITLE JOURNAL	JOURNAL REFERENCE AUTHORS TITLE JOURNAL	JOURNAL REFERENCE AUTHORS TITLE JOURNAL	COMMENT
<p>Journal of the American Medical Association 1998;279:1234-1238</p>	<p>Journal of the American Medical Association 1998;279:1234-1238</p>	<p>Journal of the American Medical Association 1998;279:1234-1238</p>	<p>Journal of the American Medical Association 1998;279:1234-1238</p>	

LOCUS AC097468 169741 bp DNA linear PRI 01-MAR-2002
DEFINITION Homo sapiens BAC clone RP11-3304 from 2, complete sequence.

KEYWORDS	SOURCE	FIG.
ORGANISM	Homo sapiens	(human)
	Homo sapiens	

REFERENCE
AUTHORS
TIMY E.
1 (bases 1 to 169741)
Sulston, J.E. and Waterston, R.
toward a complete human genome

9847074
2 (bases 1 to 169741)

AUTHORS Kruchowski, S., Cotton, M. and Doebber, A.
TITLE The sequence of Homo sapiens BAC clone RP11-3304

COMMENT

Data from AC073128, AC092152, and AC020575 was used to finish this clone, AC097468. Polymorphisms have been identified between AC073128, AC020575, and AC097468.

QY	1005	GCCTGTGTC---CAAACGCCACCGCGCGTGC	CGCCGCGCCACCGTGCCTCTACCCAAACCC	1061
Db	61437	GCCGGGCGCGGCGCGCGCCGCGCCACCGCTG	CGCCGCGCCACCGTGCCTCTACCCAGTCC	61378
QY	1062	GGGCTTCAGCCCTCCCGGCGCCCTTTGGCGG	GTGGCGGCTTCGCACTTGGGGG	1121
Db	61377	CAGCTTCAGCCCGCGCGCGCGCGCTTCGGG	CGCGTTCGCACTTGGGGG	61318
QY	1122	TCATTATCACTAGACGGACGGCGCGGTGCAG	TGGGGCGCTTCCACACAGCCAGTGA--	1179
Db	61317	CCATTACCACTAGACGGGCGGTTCGGGTGC-	CTGGCGCTTCGCCCGCACCGCTAGAGTCT	61259
QY	1180	---CCAATCCATCCTCATCTCGGAGGAGCCG	CGAAGATTTTCCCGACGTTCTTTACC	1236
Db	61258	CGCCGATCCCATCGGCATCCCGGGAGGCGCC	GGAGCGCTCGTCAACCGTCTCTAAT	61199
QY	1237	ACGATTTCTGTCAGCAGCGCTCCCGAGCCAG	GGAAGGATGGGAAGCTCTGAG	1296
Db	61198	CCAGAGTTTACTTCACCTGCGCACTTAGCAG	GGGGACG---GGACCGAAGTCTCCCTCAA	61142
QY	1297	GTCTTCTCTGAATACAGAGGCTTCAGAGCTC	CCCAATTATCATCACCCAGGAAGGTGCATG	1356
Db	61141	TCCTTGTCTGTACTAGATTGCTCTGTCTCC	ACCCCGCAGTCCCCTGAGGAGGCGATG	61082
QY	1357	TGTCGCCACTTT-AAITTTTTCTCTCCAAGT	CTCCAGATTTCTGGAATCCCGTCTTTTTT	1415
Db	61081	TGGCCCTCTTTTCACTTTTTTTCTTAGTCT	CTCCAGGTCCCGAGGGATTTGTGGACC	61022
QY	1416	TTCTCTTCTACCTGGAGCCCTGCCTCTCTTA	TATGACCCCTAGTTTTCTGTGTTGTT	1475
Db	61021	TCCTTGTCTCCGCCACACTCCAGTGCAATT	TCGCGCTGGCTCTCTAGAAGCCCAITCAAT	60962
QY	1476	TTTTTTTTTCTCTCTCTCTCTCAITTTTTT	TCCTCCACGACTACTCCAAACGCT	1535
Db	60961	ATCACTACTCTTTAAAGAGTGCCAAATCT	TTTCCCACTTTTG-TCCTCCCAAGAACT	60903
QY	1536	AGTACCTCGGTAGTACTCGAGCTTCTCACAT	CCCCCTTTTCGGGATATGAGAAGCATC	1595
Db	60902	GCTCCCACTCAGCAGTGGAGGCGCTCTCAG	GGTCTCTCTCTG-----GAC	60854
QY	1596	AAAAACATCTCTGTTGTTCATCCCTATCC	CAACACTCTGGTTCGCTCCCATTA	1655
Db	60853	CTGAGCAGGTTTGGTGAAGGCCACCGTCT	CCGTGACACACGCGCCCTTCTCTCTGTC	60794
QY	1656	CCACACTCTGGCCCAAGGACCTCGTCTGAT	ATATTCCTTTTACGCCCTTAAAGATCC	1715
Db	60793	CCACACTCCAGAGAACTCCCGGTGTGTTCT	GACCTTTTCAGCCCCCATTAAGTCTC	60734
RESULT	14			
HSDNAFEV3				
LOCUS			1403 bp	DNA
DEFINITION				linear
ACCESSION				PRI 23-APR-1997
VERSION				
KEYWORDS				
SOURCE				
ORGANISM				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
MEDLINE				
PUBMED				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
REMARK				

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 26, 2003, 12:37:51 ; Search time 492 Seconds
(without alignments)
9612.633 Million cell updates/sec

Title: US-10-027-859-1

Perfect score: 1752
Sequence: 1 ttgttttaacaacacgtt.....aaaaaaaaaaaaaaaaaa 1752

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N Geneseq 19Jun03.*
1: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.*
2: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.*
3: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.*
4: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.*
5: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.*
6: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.*
7: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.*
8: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.*
9: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.*
10: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.*
11: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.*
12: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.*
13: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.*
14: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.*
15: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.*
16: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.*
17: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.*
18: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.*
19: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.*
20: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.*
21: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.*
22: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.*
23: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.*
24: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.*
25: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1752	100.0	1752	22	AAS13672
2	1752	100.0	1752	24	ABQ25881
3	1752	100.0	1752	24	ABQ25881
C 4	318.6	18.2	723	24	ABQ252274
5	318.6	18.2	723	24	ABQ52275
C 6	262.8	15.0	723	24	ABQ52276
C 7	262.8	15.0	723	24	ABQ52277
C 8	233.4	13.3	1754	20	AAQ29137

PN	US6268216-B1.
XX	CDNA encoding Ratt
PD	Nucleotide sequenc
XX	Rat lambda73 CDNA
PF	Oligonucleotide fo
XX	Oligonucleotide fo
PR	Oligonucleotide fo
XX	Oligonucleotide fo
PA	Hypoxia-regulated

ALIGNMENTS

RESULT 1
AAS13672
ID AAS13672 standard; cDNA; 1752 BP.
XX
AC AAS13672;
XX
DT 18-DEC-2001 (first entry)
XX
DE CDNA encoding Rattus norvegicus Pet-1 transcription factor.
XX
KW Rat; EST transcription factor; Pet-1; central serotonin 5-HT neuron;
KW serotonergic receptor; human; animal; cell cycle determined disease;
KW cancer; central nervous system disorder; CNS; psychiatric disorder;
KW neurological disorder; ss.
OS Rattus norvegicus.
XX
FH Key Location/Qualifiers
CDS 112..1134
FT /*tag= a
FT /product= "Pet-1 transcription factor"

C	9	223	12.7	719	24	ABQ25880	Oligonucleotide fo
	10	223	12.7	719	24	ABQ25881	Oligonucleotide fo
	11	223	12.7	2532	22	AAQ25887	Human polynucleoti
	12	200.4	11.4	1219	25	AB268768	Nucleotide sequenc
	13	200.4	11.4	1291	25	AB268767	Nucleotide sequenc
	14	200.4	11.4	1372	25	AB268769	Nucleotide sequenc
	15	200.4	11.4	1389	25	ABX08759	Angiogenesis-associ
	16	200.4	11.4	1509	25	AB268766	Nucleotide sequenc
	17	200.4	11.4	3166	25	ABX76295	Lung cancer-associ
C	18	199.4	11.4	567	22	ABA48124	Human breast cell
	19	199.4	11.4	567	22	ABA66003	Human foetal liver
C	20	199.4	11.4	567	22	ABA33090	Probe #11556 for g
C	21	199.4	11.4	567	22	AAK14426	Human brain expres
C	22	199.4	11.4	567	22	AAK40160	Human bone marrow
C	23	199.4	11.4	567	22	AAI20932	Probe #10865 for g
C	24	199.4	11.4	567	22	AAI46176	Probe #14862 used
C	25	199.4	11.4	567	22	AAI06643	Probe #6634 used t
C	26	199.4	11.4	567	23	ABS39742	Human liver single
C	27	199.4	11.4	567	24	ABS14215	Human genome-deriv
C	28	198.4	11.3	473	22	ABA50193	Human breast cell
C	29	198.4	11.3	473	22	ABA68128	Human foetal liver
C	30	198.4	11.3	473	22	ABA35152	Probe #13618 for g
C	31	198.4	11.3	473	22	AAK16512	Human brain expres
C	32	198.4	11.3	473	22	AAK42265	Human bone marrow
C	33	198.4	11.3	473	22	AAI48339	Probe #17025 used
C	34	198.4	11.3	473	22	AAI08696	Probe #8687 used t
C	35	198.4	11.3	473	23	ABS41873	Human liver single
	36	195	11.1	1890	14	AAQ50662	Human Hum-Fli-1 ge
	37	194	11.1	2938	14	AAQ50644	Human Hum-Fli-1 ge
	38	194	11.1	2954	22	AAH02915	Human shear stress
	39	194	11.1	2957	24	ABV94304	Breast carcinoma r
	40	194	11.1	2957	24	ABK84139	Human cDNA differe
	41	190.8	10.9	1431	24	ABK73253	DNA encoding human
	42	190.6	10.9	899	23	ABL12097	Drosophila melanog
	43	180	10.3	1447	20	AAK26551	DNA encoding chick
	44	180	10.3	1528	20	AAK26552	DNA encoding chick
	45	178	10.2	2025	23	ABL05371	Drosophila melanog

Oligonucleotide fo
Oligonucleotide fo
Human polynucleoti
Nucleotide sequenc
Nucleotide sequenc
Nucleotide sequenc
Angiogenesis-assoc
Angiogenesis-assoc
Lung cancer-associ
Human breast cell
Human foetal liver
Probe #11556 for g
Human brain expres
Human bone marrow
Probe #10865 for g
Probe #14862 used t
Probe #6634 used t
Human liver single
Human genome-deriv
Human foetal liver
Probe #13618 for g
Human brain expres
Human bone marrow
Probe #17025 used t
Probe #8687 used t
Human liver single
Human Hum-Flu-1 ge
Human Hum-Flu-1 ge
Human shear stress
Breast carcinoma r
Human cDNA differe
DNA encoding human
Drosophila melanog
DNA encoding chick
DNA encoding chick
Drosophila melanog

Db 1681 TCTGTATATATCTCTTTTCAGCCCCATTAAAGATCAAGCTTCAAAAAAAAAAAAAAAAAA 1740
Qy 1741 AAAAAAAAAAAAA 1752
Db 1741 AAAAAAAAAAAAA 1752
RESULT 2
ID ABQ78589
XX ABQ78589 standard; cDNA; 1752 BP.
AC ABQ78589;
XX
XX 25-NOV-2002 (first entry)
XX Nucleotide sequence of transcription factor specific for 5-HT neurons.
XX Transcription factor; 5-HT neuron; serotonergic neuron; Pet-1;
KW ETS domain; gene; ss.
XX
XX Bacteriophage lambda 73.
XX
XX Key Location/Qualifiers
XX CDS 112..1134
XX /*tag= a
XX /product= "Pet-1"
XX
XX US2002090647-A1.
XX
XX 11-JUL-2002.
XX
XX 08-MAY-2001; 2001US-0850799.
XX
XX 27-JUL-1998; 98US-094264P.
XX 26-JUL-1999; 99US-0360779.
XX
XX (UYCA-) UNIV CASE WESTERN RESERVE.
XX
XX Deneris ES, Fyodorov DV, Hendricks TJ;
XX
XX WPI: 2002-655831/70.
XX P-PSDB; ABB78207.
XX
XX A novel composition for detecting interactive peptide in a sample, and
XX for detecting serotonergic receptor agonists or antagonists, comprises a
XX purified Pet-1 peptide
XX
XX Disclosure; Fig 1; 30pp; English.
XX
XX The present sequence encodes a transcription factor specific for 5-HT
XX (serotonergic) neurons, designated Pet-1. The Pet-1 protein contains
XX ETS domains. Compositions comprising Pet-1 are useful for detecting an
XX interactive peptide, and for identifying and testing seronergic receptor
XX agonists and antagonists.
XX
XX Sequence 1752 BP; 356 A; 613 C; 412 G; 371 T; 0 other;
Qy Query Match 100.0%; Score 1752; DB 24; Length 1752;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1752; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 TTTGTTTTTAAACAATGTTTATTAGAAAAGTAAATATTGCATAGGTCTTAGTACTTG 60
Db 1 TTTGTTTTTAAACAATGTTTATTAGAAAAGTAAATATTGCATAGGTCTTAGTACTTG 60
Qy 61 AACATCAAGTGATTCATGAACCGTGAGTATCTTCATGTAAACAGTTCTAGATGAAGAC 120
Db 61 AACATCAAGTGATTCATGAACCGTGAGTATCTTCATGTAAACAGTTCTAGATGAAGAC 120
Qy 121 CCAGGTGGCGTCTCTCTGGGGGAGAGGGTTCACGCCCCCCCCCTCAGCCCCATCCC 180
Db 121 CCAGGTGGCGTCTCTCTGGGGGAGAGGGTTCACGCCCCCCCCCTCAGCCCCATCCC 180

Qy 181 CTCACAGCTCACTCTCTCCAGTACACCGGCACCGGATGGGCTGGGATGCAGTCTCCAGGAC 240
Db 181 CTCACAGCTCACTCTCTCCAGTACACCGGCACCGGATGGGCTGGGATGCAGTCTCCAGGAC 240
Qy 241 CCCCTCCCTCTCTACACACACCTGGCTGCCGCTCCCGCCAGAGGCTTGCCGGACCCGGGG 300
Db 241 CCCCTCCCTCTCTACACACACCTGGCTGCCGCTCCCGCCAGAGGCTTGCCGGACCCGGGG 300
Qy 301 GCGTCTACTCTTCTCCCTGTCAACCCAGTCAACCGGGGGGTATCGGCACCCCAAGCGCA 360
Db 301 GCGTCTACTCTTCTCCCTGTCAACCCAGTCAACCGGGGGGTATCGGCACCCCAAGCGCA 360
Qy 361 AAGCTGACGTGCCCTCCCTCCCTGGCTGCCCTCTCCACCGCCAGTCCCGGACGG 420
Db 361 AAGCTGACGTGCCCTCCCTCCCTGGCTGCCCTCTCCACCGCCAGTCCCGGACGG 420
Qy 421 ATGACAGAGCGGCGACCTCCAGCCCTCTGCTGATCAACATGTACCTACCAAGATCCGTC 480
Db 421 ATGACAGAGCGGCGACCTCCAGCCCTCTGCTGATCAACATGTACCTACCAAGATCCGTC 480
Qy 481 GGAGATGGTCTTTTAAAGGAAGGAGAGCCGAGCTGGGGGGCGCTGAGGCCCTGCGGTA 540
Db 481 GGAGATGGTCTTTTAAAGGAAGGAGAGCCGAGCTGGGGGGCGCTGAGGCCCTGCGGTA 540
Qy 541 CAGAAAGCGAGCGGCGAGATCCAGTTGTGGCAGTTTCTACTGGAGCTGCTGGCAGACGCG 600
Db 541 CAGAAAGCGAGCGGCGAGATCCAGTTGTGGCAGTTTCTACTGGAGCTGCTGGCAGACGCG 600
Qy 601 GCGAACCGCGCTGCATCGGCTGGGAGGGGGCGCACGGCGAGTTCAAGCTCACCGACCCC 660
Db 601 GCGAACCGCGCTGCATCGGCTGGGAGGGGGCGCACGGCGAGTTCAAGCTCACCGACCCC 660
Qy 661 GACGAGGTGGCGGACGCTGGGGCGAGCGCAAGAGCAAGCCCAATATGAATACGACAAAG 720
Db 661 GACGAGGTGGCGGACGCTGGGGCGAGCGCAAGAGCAAGCCCAATATGAATACGACAAAG 720
Qy 721 CTAAGTCGAGCACTGCGCTACTACTACGACAAAAAATCATGAGCAAGGTGCACGGCAAG 780
Db 721 CTAAGTCGAGCACTGCGCTACTACTACGACAAAAAATCATGAGCAAGGTGCACGGCAAG 780
Qy 781 CGCTACCGCTACCGCTTTGACTTCACAGGGCTGCGACAGGCTTCCAGCCACCAACCCGCG 840
Db 781 CGCTACCGCTACCGCTTTGACTTCACAGGGCTGCGACAGGCTTCCAGCCACCAACCCGCG 840
Qy 841 CAGCCCCACCGCGCTGGCGCGCGCGAGCGGCGCGCGCGCGCGCGCGCGCGCGCGCTT 900
Db 841 CAGCCCCACCGCGCTGGCGCGCGCGAGCGGCGCGCGCGCGCGCGCGCGCGCGCTT 900
Qy 901 TACAAGCTCCCGCTGGTCTGGCTTCCACTGCGCTTCCCGCGGCTTCCCAACTCAACCTT 960
Db 901 TACAAGCTCCCGCTGGTCTGGCTTCCACTGCGCTTCCCGCGGCTTCCCAACTCAACCTT 960
Qy 961 ATGGCAGCTCGGCGCGCTGGGCGCGCGCTGCTTCTTACTGGCTGTGTCCTCAACGCC 1020
Db 961 ATGGCAGCTCGGCGCGCTGGGCGCGCGCTGCTTCTTACTGGCTGTGTCCTCAACGCC 1020
Qy 1021 ACCGCGCTCGCGCGCGCGCGCTGCGCTTACCCACCGCGGCTTGCAGCGCCCTCCC 1080
Db 1021 ACCGCGCTCGCGCGCGCGCGCTGCGCTTACCCACCGCGGCTTGCAGCGCCCTCCC 1080
Qy 1081 GGGCCCTTTTGGCGGCTGGCGCGCTTCCGACCTTGGGGGGGTCAATATCACTAGACGGGA 1140
Db 1081 GGGCCCTTTTGGCGGCTGGCGCGCTTCCGACCTTGGGGGGGTCAATATCACTAGACGGGA 1140
Qy 1141 CGGCGGGGTGAGTGGGGCTCTCCACACAGCGAGTGACCAATCCCATCTCATCTCTGG 1200
Db 1141 CGGCGGGGTGAGTGGGGCTCTCTCCACACAGCGAGTGACCAATCCCATCTCATCTCTGG 1200
Qy 1201 GAGGAGCCCGAGAGTTTCCCGGACGTTCTTTTACACAGATTTGTTGACGAGCGGCT 1260
Db 1201 GAGGAGCCCGAGAGTTTCCCGGACGTTCTCTTTTACACAGATTTGTTGACGAGCGGCT 1260
Qy 1261 CCAGGCCAGGGAAGAAAGATGGGAAGCCCTCTGAGGTCTTCTCTTGAATACAGGCTTCC 1320


```

QY 646 AAGCTCACGACCCGACGAGGTGGCGGAGCTGGGGGAGCGCAAGAGCAGCCCAAT 705
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 587 AACTACGACCGGACGAGGATACGCGAGGATATAAAGCAACGCAAAACCAACCAAC 528
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 706 ATGAATACGACAAAGTAAAGTCAGCACTGCGTACTACTACGACAAACATCATGAGC 765
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 527 ATAACTACGACAAACTAAACCGCGCTACGCTACTACTACGACAAACATCATAAAC 468
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 766 AAGGTACGCGGACGCTACGCTACGCTTTGACTTCCAGGCTGGCAGGCTTGC 825
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 467 AAAATACATAAACAACGCTACGCTACGCTTCGACTTCCAAACCTAACGCAAACTAC 408
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 826 CAGCCACACCCCGGACGCGCCGCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCG 885
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 407 CAACCGCGCGCGCGCGCGCTACATACGCGCGCGCGCGCGCGCGCGCGCGCGCG 348
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 886 CAGGATGGCGCACTTAAAGCTCCGCGTGTCTGGCTCCACTGCGCTTCCCGCGCTC 945
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 347 CAAAACGACGCGCTTACAAACTACCGCGCGCGCTGCGCGCGCTACCGTTCGCGACCTC 288
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 946 TCCAACTCAACCTTATGCGAGCTGCGCGCGCGCTGCGCGCGCTGCGCTTCTTACTGG 1005
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 287 TCCAACTCAACCTTATGCGAGCTGCGCGCGCGCTGCGCGCGCTGCGCTTCTTACTAA 228
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1006 CTGGTCCCAA---CGCCACCGCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCG 1062
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 227 CCGAACCGCAACCCCGCGCGCGCTACCGCGCGCGCGCGCGCGCGCTTACCCCAATCC 168
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1063 GGTTCGACGCGCTTCCCGCGCGCTTGGCGCGCTTGGCGCGCTTGGCGCGCTTGGCGG 1122
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 167 AACTTCAACCGCGCGCGCGCGCTTGGAAACCGTAAACCGCAACCTCGCACTTAAAAAC 108
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1123 CATTATCACTAGACGCGCGCGCGCTGCG 1151
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 107 CATTACCACTAAACGAAACGATCGAATAC 79

```

```

RESULT 5
ABQ52275
ID ABQ52275 standard; DNA; 723 BP.
XX ABQ52275;
AC ABQ52275;
XX
DT 12-JUL-2002 (first entry)
XX
DE Oligonucleotide for detecting cytosine methylation SEQ ID NO 38866.
XX
KW Human; cytosine methylation; 5'-CpG-3'; uracil; cytosine; diagnosis;
KW drug; side effect; cancer; central nervous system; cardiovascular;
KW gastrointestinal; respiratory system; single nucleotide polymorphism;
KW SNP; cell differentiation; ds.
XX
OS Homo sapiens.
XX
PN WO200218632-A2.
XX
PD 07-MAR-2002.
XX
PF 01-SEP-2001; 2001WO-BP10074.
XX
PR 01-SEP-2000; 2000DE-1043826.
XX
PR 05-SEP-2000; 2000DE-1044543.
XX
PA (EPiG-) EPIGENOMICS AG.
XX
PI Olek A, Piepenbrock C, Berlin K, Guetig D;
XX
DR WPI; 2002-371829/40.
XX
PT Determining the degree of cytosine methylation in genomic DNA, useful
PT for diagnosis and prognosis, comprises selective hybridization of
PT amplicons from chemically treated DNA

```

```

PS Claim 12; 56pp + Sequence Listing; 56pp; German.
XX
CC This invention describes a novel method for determining the degree of
CC methylation of a particular cytosine in a motif 5'-CpG-3', present in a
CC genomic sample of DNA. The sample is treated chemically to convert
CC cytosine (C) but not methylated C to uracil, then part of the genomic
CC DNA that contains the target C is amplified to form a labeled amplicon.
CC The amplicon is hybridised to two classes, each with at least one
CC member, of oligonucleotides and/or peptide-nucleic acid (PNA) oligomers
CC and the degree of hybridisation to both classes is determined from the
CC label on the amplicon. From the ratio of labels hybridised to the two
CC classes of oligomers, the degree of methylation is calculated. The method
CC is used: (i) for diagnosis and/or prognosis of side effects of
CC therapeutic drugs and of a wide range of diseases, e.g. cancer, disorders
CC of the central nervous, cardiovascular, gastrointestinal and respiratory
CC systems etc., particularly by detecting mutations or single nucleotide
CC polymorphisms (SNP's); and (ii) for differentiation of cell or tissue
CC types and for investigating cell differentiation. The method allows the
CC methylation status of many C residues to be determined simultaneously.
CC ABQ3410-ABQ54121 represent genomic DNA sequences used to illustrate the
CC method for determining the degree of cytosine methylation described in
CC the disclosure of the invention.
XX
SQ Sequence 723 BP; 239 A; 301 C; 89 G; 94 T; 0 other;
Query Match 18.2%; Score 318.6; DB 24; Length 723;
Best Local Similarity 70.3%; Pred. No. 9.7e-54;
Matches 442; Conservative 0; Mismatches 184; Indels 3; Gaps 1;
QY 526 CTGAGCCCTGCGGTACAGAAAGCGGAGCGGAGTCCAGTCTTGGCAGTTCTTACTGGAG 585
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 586 CTGCTGGCAGACCGCGCGAAACCGCGCTGCGTGGGGGCGCGCGCGCGAGTTTC 645
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 646 AAGCTCACGACCGCGCGAGGTGGCGGAGCGCTGGGGGCGCGCGCGAGGAGCGCAAT 705
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 137 AAATCTACGAAACCGCGCGAAACGAAATACGCGAGTAAACGAAACGAAACCAAC 196
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 706 ATGAATACGACAAAGTAAAGTCAGCACTGCGCTACTACTACGACAAACATCATGAGC 765
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 197 ATAACTACGACAAACTAAACCGCGCTTACTACTACTACGACAAACATCATAAAC 256
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 766 AAGGTGACGCGCAAGCGCTACGCTTACCGCTTTGACTTCCAGGCTGGCAGCGCTTGC 825
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 257 AAATACATATAACAAACGCTACGCTTACCGTTTCGACTTCCAAACCTTACGCAAACTAC 316
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 826 CAGCCACACCGCGCGAGCGCGCGCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCG 885
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 317 CAACCGCGCGCGCGCGCTCATACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 376
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 886 CAGGATGGCGCACTTACAAAGCTCCCGCTGCTGGCTTCCACTGCCCTTCCCGCGCTC 945
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 377 CAACGACGCGCTTACAAACTACCGCGCGCGCTTCCCGCGCTTCCCGCGCTC 436
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 946 TCCAACTCAACCTTATGCGAGCTTGGCGCGCGCGCTGCGCGCGCTTCTTACTGG 1005
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 437 TCCAACTCAACCTTATGCGAGCTTCCGCGCGCGCGCGCGCGCGCGCGCGCGCG 496
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1006 CTGGTCCCAA---CGCCACCGCGCTGCGCGCGCGCGCGCGCGCTTACCCCAACCG 1062
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 497 CCGAACCGCGCGCGCGCGCTACCGCTACCGCGCGCGCGCGCGCGCGCGCGCGCG 556
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1063 GGTTCGACGCGCTTCCCGCGCGCTTGGCGCGCTTGGCGCGCTTGGCGCGCTTGGCGG 1122
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 557 AACTTACAAACCGCGCGCGCGCTTCCGAAACCGTAAACGCGCAACCTCGCACTTAAAAAC 616
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1123 CATTATCACTAGACGCGCGCGCGCTGCG 1151
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 617 CATTACCACTAAACGAAACGATCGAATAC 645

```


XX Claim 12; 56pp + Sequence Listing; 56pp; German.

XX This invention describes a novel method for determining the degree of
 CC methylation of a particular cytosine in a motif 5'-CpG-3', present in a
 CC genomic sample of DNA. The sample is treated chemically to convert
 CC cytosine (C) but not methylated C, to uracil, then part of the genomic
 CC DNA that contains the target C is amplified to form a labeled amplicon.
 CC The amplicon is hybridised to two classes, each with at least one
 CC member, of oligonucleotides and/or peptide-nucleic acid (PNA) oligomers
 CC and the degree of hybridisation to both classes is determined from the
 CC label on the amplicon. From the ratio of labels hybridised to the two
 CC classes of oligomers, the degree of methylation is calculated. The method
 CC is used: (i) for diagnosis and/or prognosis of side effects of
 CC therapeutic drugs and of a wide range of diseases, e.g. cancer, disorders
 CC of the central nervous, cardiovascular, gastrointestinal and respiratory
 CC systems etc., particularly by detecting mutations or single nucleotide
 CC polymorphisms (SNP's); and (ii) for differentiation of cell or tissue
 CC types and for investigating cell differentiation. The method allows the
 CC methylation status of many C residues to be determined simultaneously.
 CC ABQ13410-ABQ54121 represent genomic DNA sequences used to illustrate the
 CC method for determining the degree of cytosine methylation described in
 CC the disclosure of the invention.

XX Sequence 723 BP; 306 A; 229 C; 89 G; 99 T; 0 other;

Query Match 15.0%; Score 262.8; DB 24; Length 723;
 Best Local Similarity 65.8%; Pred. No. 1.1e-42;
 Matches 400; Conservative 0; Mismatches 207; Indels 3; Gaps 1;

QY 544 AAGGCGGCGGCGAGATGTCGAGTTGTGGCAGTTTCTACTGAGCTGTGGCAGACCGCGG 603
 DB 689 ATAGGTAGCGGATAGATTAGTTGTGGTAGTTTGTGGAGTTGTGGTAGTCGCGG 630
 QY 604 AACGCCGTGTGATCCGTGGGAGCGCGCCAGCGAGTTCAAGCTCACCGACCCCGAC 663
 DB 629 AACGTCGGTTGTATCGTGGGAGCGGTTACGGCGAGTTTAAAGTTTACGGATTCCGAC 570
 QY 664 GAGGTGGCGGCGCTGCGGCGGAGCGCAAGAGCAAGCCCAATATGAACTACGACAGCTA 723
 DB 569 GAGGTGGCGGCGGTTGGGCGGAGCGTAAGAGTAAGTTTAAATATGAATACGATAAGTTG 510
 QY 724 AGTCGAGCAGCTGCGCTACTACTACGACAAACATCATGACGAGGTGACCGCAGCGC 783
 DB 509 AGTCGCGTTTGGGTTATTATACGATAGATATATATGATAGGTGTATGTAAGCGT 450
 QY 784 TAGCCCTACCGTTTGCATTTCCAGGCGCTGGCACAGCTTGCACGACCCCGCGCAC 843
 DB 449 TAGCTTTATCGTTTCGATTTTAAAGGTTTGGCGTAGGTTTGTAGTCGTTCGCGTAC 390
 QY 844 GCCCAGCGCGCTGCGCGCGCGGAGCGGAGCGCGCCAGGATGCGGACCTTTAC 903
 DB 389 GTTATGTCGTCGTCGAGTTGTGTCGTCGTCGTCGTTTGGAGCGCGCGGTTTAT 330
 QY 904 AAGCTCCGCGCTGCTGGCTCCACTGCGCTTCCCGGCGCTCTCCAAACTCAACCTTATG 963
 DB 329 AAGTTGTCGTCGTTTCTGTTGTTTCTGTTTCTGTTTCTGTTTCTGTTTCTGTTTCTG 270
 QY 964 GCAGCCTCGCGCGCTGCGCGCGCTGCTGCTTCTTACTGCGCTGCTCC---CAACGCC 1020
 DB 269 GTCGTTTCGCTCGGCTGCGGTTTCGTCGCTTTTATTTGTCGCGGTTTCGCTTCGTC 210
 QY 1021 ACCGCGCGCTGCGCGCGCGCTGCTGCTTACCCCAACCCCGGCTTGCAGCCCTCC 1080
 DB 209 GTTATGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTC 150
 QY 1081 GGGCCCTTTGGCGCGGTGGCGCGCTGCTGCTGCTGCGGCTGCTTATCATCTAGACGGA 1140
 DB 149 GGGTTTTTCGGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 90
 QY 1141 CGCGCGGCTG 1150
 DB 89 CGGTCGGGTG 80

RESULT 8

AAAX29137/c
 ID AAX29137 standard; cDNA; 1754 BP.
 AC AAX29137;
 XX
 DT 04-JUN-1999 (first entry)
 XX
 DE Hypoxia-regulated gene sequence RTP801.
 XX
 KW Hypoxia-regulated gene; therapeutic; diagnostic; hypoxia; ischemia;
 KW apoptosis; angiogenesis; tumorigenic cell; trauma; limb reattachment;
 KW revascularisation; ss.
 XX
 OS Rattus sp.
 XX
 PN WO9909049-A1.
 XX
 PD 25-FEB-1999.
 XX
 PF 21-AUG-1998; 98WO-US17296.
 XX
 PR 21-AUG-1997; 97US-0056453.
 XX
 PA (KOHN/) KOHN K I.
 PA (QUAR-) QUARK BIOTECH INC.
 XX
 PI Binat P, Skaliter R;
 XX
 DR WPI; 1999-180965/15.
 DR P-PSDB; AAY03634.
 XX
 PT New isolated hypoxia-related genes - used to develop products for
 PT use in therapy and diagnosis in e.g. hypoxia, ischemia, apoptosis
 PT and angiogenesis
 XX
 PS Claim 1; Page 69-70; 92pp; English.
 XX
 CC Sequences AAX29137 to AAX29142 represent isolated hypoxia-regulated
 CC genes. The genes and their products can be therapeutically and
 CC diagnostically used in hypoxia, ischemia, apoptosis and angiogenesis. The
 CC products and methods can be used for e.g. inducing apoptosis in
 CC tumorigenic cells or angiogenesis in trauma situations where e.g. a limb
 CC must be reattached or in a transplant where revascularisation is needed.
 XX
 SQ Sequence 1754 BP; 356 A; 459 C; 546 G; 393 T; 0 other;

Query Match 13.3%; Score 233.4; DB 20; Length 1754;
 Best Local Similarity 99.6%; Pred. No. 8.6e-37;
 Matches 234; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TTTGTTTAAACAAACATGTTTATTAGAAAAGTAAAAATATTCATAGGCTTAGTACTTG 60
 DB 1733 TTTGTTTAAACAAACATGTTTATTAGAAAAGTAAAAATATTCATAGGCTTAGTACTTG 1674
 QY 61 AACATCAAGTGATTTTCATGAACCGTGAGTATCTTCATGTAAACAGTTCTAGATGGAAGAC 120
 DB 1673 AACATCAAGTGATTTTCATGAACCGTGAGTATCTTCATGTAAACAGTTCTAGATGGAAGAC 1614
 QY 121 CCAGGTGGCGTCTCTCTGGGGAGAGGGTTCCAGCCCGCCCGCCCTCAGCCCGCATCC 180
 DB 1613 CCAGGTGGCGTCTCTCTGGGGAGAGGGTTCCAGCCCGCCCGCCCTCAGCCCGCATCC 1554
 QY 181 CTCAGAGTCACTCTCTCCAGTACACCGGATGGGATGGGATGGGATGGGATGGGATGGG 235
 DB 1553 CTCAGAGTCACTCTCTCCAGTACACCGGATGGGATGGGATGGGATGGGATGGGATGGG 1499

RESULT 9

ABQ25880/c
 ID ABQ25880 standard; DNA; 719 BP.

XX AC ABQ25880;
XX DT 12-JUL-2002 (first entry)
XX DE Oligonucleotide for detecting cytosine methylation SEQ ID NO 12471.
XX KW Human; cytosine methylation; 5'-CpG-3'; uracil; cytosine; diagnosis;
XX KW drug; side effect; cancer; central nervous system; cardiovascular;
XX KW gastrointestinal; respiratory system; single nucleotide polymorphism;
XX KW SNP; cell differentiation; ds.
XX OS Homo sapiens.
XX PN WO200218632-A2.
XX PD 07-MAR-2002.
XX PF 01-SEP-2001; 2001WO-EP10074.
XX PR 01-SEP-2000; 2000DE-1043826.
XX PR 05-SEP-2000; 2000DE-1044543.
XX PA (EPiG-) EPIGENOMICS AG.
XX PI Olek A, Piepenbrock C, Berlin K, Guetig D;
XX DR WPI; 2002-371829/40.
XX PT Determining the degree of cytosine methylation in genomic DNA, useful
XX PT for diagnosis and prognosis, comprises selective hybridization of
XX PT amplicons from chemically treated DNA -
PS Claim 12; 56pp + Sequence Listing; 56pp; German.
CC This invention describes a novel method for determining the degree of
CC methylation of a particular cytosine in a motif 5'-CpG-3', present in a
CC genomic sample of DNA. The sample is treated chemically to convert
CC cytosine (C) but not methylated C, to uracil, then part of the genomic
CC DNA that contains the target C is amplified to form a labeled amplicon.
CC The amplicon is hybridised to two classes, each with at least one
CC member, of oligonucleotides and/or peptide-nucleic acid (PNA) oligomers
CC and the degree of hybridisation to both classes is determined from the
CC label on the amplicon. From the ratio of labels hybridised to the two
CC classes of oligomers, the degree of methylation is calculated. The method
CC is used: (i) for diagnosis and/or prognosis of side effects of
CC therapeutic drugs and of a wide range of diseases, e.g. cancer, disorders
CC of the central nervous, cardiovascular, gastrointestinal and respiratory
CC systems etc., particularly by detecting mutations or single nucleotide
CC polymorphisms (SNP's); and (ii) for differentiation of cell or tissue
CC types and for investigating cell differentiation. The method allows the
CC methylation status of many C residues to be determined simultaneously.
CC ABQ13410-ABQ54121 represent genomic DNA sequences used to illustrate the
CC method for determining the degree of cytosine methylation described in
CC the disclosure of the invention.
XX SQ Sequence 719 BP; 113 A; 83 C; 268 G; 255 T; 0 other;
Query Match 12.7%; Score 223; DB 24; Length 719;
Best Local Similarity 61.2%; Pred. No. 8e-35;
Matches 413; Conservative 0; Mismatches 255; Indels 7; Gaps 3;
QY 609 CGGCTGCTCGCTGGGAGGGCGCCACGGCGAGTTCAAGTTCACCGACCGGAGGT 668
DB 711 CGACTACATCGGTAAAAAAGCATCAGCGAGATTCAAACTCAGAACCCGACGAAAT 652
QY 669 GGGCGGAGCTGGGCGAGCGGAGGAGCAAGCCCAATATGACTACGACAGCTAAGTCG 728
DB 651 AACGCGAGTAAACAGCAACGCAAAACCAACCAATTAACCTACGACAACTAAACCG 592
QY 729 AGCACTCGGCTACTACTACGCAAAACATCATGAGCAAGGTGACGCGCAAGCGCTACGC 788
DB 591 CGCCCTACTAT-CTACTACGACAAACATCATTAACCAATATACATAATAACGCTACGC 533

QY 789 CTACGCTTTGACTTTCAGGGCTGGCACAGGCTTGCCAGCGCACCAACCGCGCAGCCCA 848
DB 532 CTACGCTTTCAGTTCCTCAAAACCTAACGCAATCTACCAACCGCGCGCGCGCTCA 473
QY 849 CGCGCGCTGCTGCGCGCGCGAGCGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 908
DB 472 TACGCGCGCGCGCAACTACTACGCGCGCGG-CGTGCGCGCGCGCGCGCGCGCGCG 414
QY 909 CCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 968
DB 413 ACCGCGCGCGCTGCGACCGCTACCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 354
QY 969 CTGCGCGCGCGCTGCGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1028
DB 353 CTGCTAAATAATCGCGCGCGAGGACTTCTCTACTAAACCGACCGGATACCGAC 294
QY 1029 TCGCGCGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1088
DB 293 TAAACCGCGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 234
QY 1089 TGGCGCGGTGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1148
DB 233 CGACCGTAAACGCTCAACCTGCGGCTTAAACCAACCTACCACTAAACCAACCA 174
QY 1149 TGC-----AGTGGGGCGCTCTCCACACAGCGAGTGAACCAATCCCTCATCT 1203
DB 173 TACCTACGACCTCGACCGAGCGTCTCTAAATCTCGCGCTATCCCATCTACATCT 114
QY 1204 GAGCGCGGAGATTTCCCGGAGCGTCTCTTTTACCACAGATTTGGTTGCGAGCG 1263
DB 113 AATCCGAAACCTCCGTCACCGCTCTCTAAATCCAAATTTACTCCACCTACCGCA 54
QY 1264 AGCCGAGGAGAGAA 1278
DB 53 AACAAAAACGAA 39
RESULT 10
ABQ25881
ID ABQ25881 standard; DNA; 719 BP.
XX AC ABQ25881;
XX DT 12-JUL-2002 (first entry)
XX DE Oligonucleotide for detecting cytosine methylation SEQ ID NO 12472.
XX KW Human; cytosine methylation; 5'-CpG-3'; uracil; cytosine; diagnosis;
XX KW drug; side effect; cancer; central nervous system; cardiovascular;
XX KW gastrointestinal; respiratory system; single nucleotide polymorphism;
XX KW SNP; cell differentiation; ds.
XX OS Homo sapiens.
XX PN WO200218632-A2.
XX PD 07-MAR-2002.
XX PF 01-SEP-2001; 2001WO-EP10074.
XX PR 01-SEP-2000; 2000DE-1043826.
XX PR 05-SEP-2000; 2000DE-1044543.
XX PA (EPiG-) EPIGENOMICS AG.
XX PI Olek A, Piepenbrock C, Berlin K, Guetig D;
XX DR WPI; 2002-371829/40.
XX PT Determining the degree of cytosine methylation in genomic DNA, useful
XX PT for diagnosis and prognosis, comprises selective hybridization of
XX PT amplicons from chemically treated DNA -

XX	Claim 12; 56pp + Sequence Listing; 56pp; German.
XX	This invention describes a novel method for determining the degree of
XX	methylation of a particular cytosine in a motif 5'-CpG-3', present in a
CC	genomic sample of DNA. The sample is treated chemically to convert
CC	cytosine (C) but not methylated C to uracil, then part of the genomic
CC	DNA that contains the target C is amplified to form a labeled amplicon.
CC	The amplicon is hybridised to two classes, each with at least one
CC	member, of oligonucleotides and/or peptide-nucleic acid (PNA) oligomers
CC	and the degree of hybridisation to both classes is determined from the
CC	label on the amplicon. From the ratio of labels hybridised to the two
CC	classes of oligomers, the degree of methylation is calculated. The method
CC	is used: (i) for diagnosis and/or prognosis of side effects of
CC	therapeutic drugs and of a wide range of diseases, e.g. cancer, disorders
CC	of the central nervous, cardiovascular, gastrointestinal and respiratory
CC	systems etc., particularly by detecting mutations or single nucleotide
CC	polymorphisms (SNPs); and (ii) for differentiation of cell or tissue
CC	types and for investigating cell differentiation. The method allows the
CC	methylation status of many C residues to be determined simultaneously.
CC	ABQ13410-ABQ54121 represent genomic DNA sequences used to illustrate the
CC	method for determining the degree of cytosine methylation described in
CC	the disclosure of the invention.
XX	
SQ	Sequence 719 BP; 255 A; 268 C; 83 G; 113 T; 0 other;
	Query Match 12.7%; Score 223; DB 24; Length 719;
	Best Local Similarity 61.2%; Pred. No. 8e-35;
	Matches 413; Conservative 0; Mismatches 255; Indels 7; Gaps 3
Qy	609 CGGTGCATCGCTGGAGGCGGCCACGGCGAGTTCAAGTCCAGCACCGACGAGGT 668
Db	9 CGACTATCGCTTAATAAAAAACGATCAGACGAATTCAAATTCAGAAACCCGAACAAAT 68
Qy	669 GGCGCGAGCTGGGCGGAGCCAGAGACAAGCCCAATATGACTACGACAGCTAAGTCG 728
Db	69 AACCGGAGGATAAACGAGAACGCAAAAACAACCACATATAACTACGACAACTAAACCG 128
Qy	729 AGCACTCGCTACTACTACGACAAAAACATCATGACGAAGGTGCAAGCGCTACGC 788
Db	129 CGCCCTACTAT-CTACTACGACAAAAACATCATRAACAAAAATACATAAACGCTACGC 187
Qy	789 CTACCGCTTTGACTTCCAGGCGCTGGCAGAGCTTGCCAGGCACACCCGGCGACGCCCA 848
Db	188 CTACCGCTTGACTTCCAACAACTTAAACGCAATCTACCAACGGCGCGCGCAGCGCTCA 247
Qy	849 CGCGCGCGCTCGCGCGCGCAGCGGACGCGCGCCCGCAGATGCGCACTTTACAAGCT 908
Db	248 TACCGCCGCGCACTACTACGCGCGCG-CGTGCCCAAAACGACGGCTCTACAACT 306
Qy	909 CCGCGCTGTGTGTGCTTCACTTGCCCTTCCCAGGCTCTTCCAAACTCAACCTTATGGCAGC 968
Db	307 ACCCGCGACCTCGACCGCGCTACCGCTTCCGCGACCTTCCAACCTCAACCTCATAACCGC 366
Qy	969 CTCGGCGCGGTGCGCGCGCTGGCTTCTCTTACTTGGCTCTGTCCTCAACGCCACCGCGCG 1026
Db	367 CTGCTAAAAATCGCGCGACGACTTCTCTTACTAACCGAACGNATACGACGCCACCGC 426
Qy	1029 TGCGCGCGCCACCGCTCGCTCTACCCAAACCGGGCTTGAGCGCCCTTCCGGGCGCCCTT 1088
Db	427 TAAAAACGCCAAACGCGCGCTCTAAACCCAACTCGAACTTCAACCCCAACACGAAACATT 486
Qy	1089 TGGCGCGGTGCGCGCGCTTCGACTTTGGGGGTTCAATATCACTAGACGGACGCGCGCG 1148
Db	487 CGACCGGTAAACGTTAAACCTCGACTTTAAAAAACATTTACCACCTAAACAAAACGATCGAA 546
Qy	1149 TGC-----AGTGGGGCTCTTCCCAACAGCCAGTGACCAATCCCATCTCATCTCTGGGAG 1203
Db	547 TACCTACGACCTCGACCGAACGTCCTAAAAATCTCGCGTATCCCATCTACATCTCGAAA 606
Qy	1204 GAGCCCGAAGATTTCCCGGACGTTTCCCTTTACACAGATTTGTTGTGACGAGCGCGCTCCC 1263
Db	607 AATCCCGANAACTCCGCTCAACCGCTCTCTAATCCAAATTTACTCAACTTACCGCACTT 666

Query Match 12.7%; Score 223; DB 22; Length 2532;
Best Local Similarity 72.2%; Pred. No. 1.1e-34;
Matches 366; Conservative 0; Mismatches 70; Indels 71; Gaps 3;

Qy 546 AGGAGGGGCGAGATCCAGTTGTGGCAGTTTCTACTGGAGTGTGGCAGACCGCGGAA 605
Db 2096 ATGACGGGAGAGATCCAGCTGTGGCAGTTTCTCTGGAGTGTGGCAGACCGCGGAA 2155

Qy 606 CGCGGCTGCAATCCGTGGAGGCGCGCCACGGGAGTTCAAGCTCACCGACCCCGACGA 665
Db 2156 CG-CGGCTGCATCCGTGGAGGCGGTCAAGGAGTTCAAGCTCACCGACCCCGACGA 2214

Qy 666 GGTGGCGGAGCGTGGGCGGAGCGCAAGAGCAAGCCCAATATGAATACGACAAAGCTAAG 725
Db 2215 GGTGGCGGCGGTGGGCGGAGCGCAAGAGCAAGCCCAATGAATACGACAAAGCTGAG 2274

Qy 726 TCGAGCACTGGCTACTACTACGACAAACATCATGACGAGGTGACGCAAGCGCTA 785
Db 2275 CCGCGCCCTGCTAT-CTACTACGACAAACATCATGACGAGGTGATGTAAGCGCTA 2333

Qy 786 CGCCTACCGCTTTGACTTCCAGGGCGCTGGCAGAGGCTTGCCAGCCACCACCGCGCAGC 845
Db 2334 CGCCTACCGCTTGCATTTCCAGGGCGCTG----- 2361

Qy 846 CCACGCGCGCGCTGCGCGCGCGGAGCGGAGCGCGCGCGCGCGCGCGCGCGCGCGCG 905
Db 2362 -----GCGCAGGACGCGCGGCTCTACAA 2384

Qy 906 GCTCCCGGCTGGTGGCTCCACTGCGCTTCCCGGCGCTCTCCAACTCAACCTTATGGC 965
Db 2385 GCTCCCGGCGCGCTGACACCGCGCTTCCCGGCGCTCTCCAACTCAACCTTATGGC 2444

Qy 966 AGCTCGCGCGCGCTGGGCGCGCGCTGCTTCTTACTGGCGCTGTGCCAACGCGCACCGC 1025
Db 2445 CGCTTCGCTGAGTTCGCGCGCGCGCTTCTTACTGGCGCGCGCGCGCGCGCGCGCGC 2504

Qy 1026 CGTTCGCGCGCGCGCGCGCGCTGCTTCTA 1052
Db 2505 CGCTGAAGCGCGCGCGCGCGCGCTTCTA 2531

RESULT 12

ABZ68768
ID ABZ68768 standard; DNA; 1219 BP.

XX AC ABZ68768;

XX DT 16-MAY-2003 (first entry)

XX DE Nucleotide sequence of human ERG splice variant C-1-3.

XX KW Human; splice variant; ERG; Ets; transcription factor; C-1-1; C-1-2;
C-1-3; C-1-4; chondrocyte; articular chondrocyte; bone disease;
cartilage disease; tumour; gene; ss.

XX OS Homo sapiens.

XX FH Key Location/Qualifiers

XX FT CDS 146..1156

XX FT /*tag= a

XX FT /product= "ERG splice variant C-1-3"

XX PN WO2003000724-A2.

XX PD 03-JAN-2003.

XX PF 08-MAR-2002; 2002WO-EP02605.

XX PR 25-JUN-2001; 2001US-300756P.

XX PA (TIGSE-) TIGENIX NV.

XX PI Luyten F, De Bari C, Dell'Accio F;

XX DR

WPI; 2003-221398/21.

DR P-PSDB; ABP97697.

XX PT

Novel splice variants of the Ets-related gene (ERG) transcription factor, useful for preventing or treating bone-related or cartilage-related diseases or Ets-related gene associated tumour

XX PS

Claim 11; Fig 6; 60pp; English.

XX CC

The present sequence encodes splice variant C-1-3 of human ERG. ERG belongs to the Ets family of transcription factors. The specification describes C-1-1 (characterised by the absence of the 81 bp fragment, and the presence of the 218 bp and 72 bp fragments of the ERG-3 cDNA sequence); C-1-2 (characterised by the absence of the 81 bp and 218 bp fragments, and the presence of the 72 bp fragment of the ERG-3 cDNA sequence); C-1-3 (characterised by the absence of the 72 bp, 81 bp and 218 bp fragments of the ERG-3 cDNA sequence); and C-1-4 (characterised by the absence of the 218 bp fragment, and the presence of the 72 bp and 81 bp fragments of the ERG-3 cDNA sequence). The splice variant or cells with a chondrocytes-like phenotype, for preventing de-differentiation of articular chondrocytes during in vitro culture expansion, for preventing or treating bone-related or cartilage-related diseases or Ets-related gene associated tumour, or as marker of progenitor cells of stable chondrocytes. Cells expressing the polynucleotide are useful for producing or repairing cartilage tissue in a mammal.

XX SQ

Sequence 1219 BP; 312 A; 393 C; 288 G; 226 T; 0 other;

Query Match

Best Local Similarity 11.4%; Score 200.4; DB 25; Length 1219;

Matches 243; Conservative 0; Mismatches 71; Indels 0; Gaps 0;

Qy 529 AGCCCTCGGTACAGAAAGGCGGCGAGATCCAGTTGTGGCAGTTTCTACTGGAGCTG 588

Db 617 AGCCGCTTGCAATCCAGGCGAGTGGCCAGATCCAGCTTTGGCAGTTTCTCTCGAGCTC 676

Qy 589 CTGGCAGACCGCGGAGCGCGGCTGCTCGTGGGAGGCGCGCGGAGTTTCAAG 648

Db 677 CTGTGCGACAGCTCCAACTCCAGCTGCATCACCTGGGAAGGCACCAACGGGAGTTCAAG 736

Qy 649 CTACCCGACCCCGAGAGGTGGCGGCGAGCGCTGGGCGGCGCAAGAGCAAGCCCAATATG 708

Db 737 ATGACGGATCCGACGAGGTGGCCCGGCTGGGAGAGCGAGAGCAAAACCCCAATG 796

Qy 709 AACTACGACAAAGCTAAGTCGAGCACTCGCTACTACTACGACAAACCATCATGAGCAAG 768

Db 797 AACTACGATAAGCTCAGCCGCGCTCCGTTACTATGACAAAGACATCATGACCAAG 856

Qy 769 GTGCAGCGCAAGCGCTACGCTACCGCTTTGACTTCCAGGCGCTGGCAGAGCTTGCCAG 828

Db 857 GTCCATGGGAAGCGCTACGCTACAAGTTCCAGCTTCCACGGGATCGCCCGGCTCCAG 916

Qy 829 CCACCACCCCGCA 842

Db 917 CCCACCCCGCA 930

RESULT 13

ABZ68767

ID ABZ68767 standard; DNA; 1291 BP.

XX AC ABZ68767;

XX DT 16-MAY-2003 (first entry)

XX DE Nucleotide sequence of human ERG splice variant C-1-2.

XX KW Human; splice variant; ERG; Ets; transcription factor; C-1-1; C-1-2;
C-1-3; C-1-4; chondrocyte; articular chondrocyte; bone disease;
cartilage disease; tumour; gene; ss.

```

XX OS Homo sapiens.
XX FH Key Location/Qualifiers
XX FT CDS 146..1228
XX FT FT /*tag= a
XX FT FT /product= "ERG splice variant C-1-2"
XX PN WO2003000724-A2.
XX PD 03-JAN-2003.
XX PF 08-MAR-2002; 2002WO-EP02605.
XX PR 25-JUN-2001; 2001US-300756P.
XX PA (TIGE-) TIGENIX NV.
XX PI Luyten F, De Bari C, Dell'Accio F;
XX DR WPI: 2003-221398/21.
XX DR P-PSDB; ABP97696.
XX PT Novel splice variants of the Ets-related gene (ERG) transcription
XX PT factor , useful for preventing or treating bone-related or
XX PT cartilage-related diseases or Ets-related gene associated tumour
XX PS Claim 10; Fig 5; 60pp; English.
XX CC The present sequence encodes splice variant C-1-2 of human ERG. ERG
XX CC belongs to the Ets family of transcription factors. The specification
XX CC describes C-1-1 (characterised by the absence of the 81 bp fragment, and
XX CC the presence of the 218 bp and 72 bp fragments of the ERG-3 cDNA
XX CC sequence); C-1-2 (characterised by the absence of the 81 bp and 218 bp
XX CC fragments, and the presence of the 72 bp fragment of the ERG-3 cDNA
XX CC sequence); C-1-3 (characterised by the absence of the 72 bp, 81 bp and
XX CC 218 bp fragments of the ERG-3 cDNA sequence); and C-1-4 (characterised by
XX CC the absence of the 218 bp fragment, and the presence of the 72 bp and
XX CC 81 bp fragments of the ERG-3 cDNA sequence). The splice variant
XX CC polynucleotides are useful for stabilizing the phenotype of chondrocytes
XX CC or cells with a chondrocytes-like phenotype, for preventing
XX CC de-differentiation of articular chondrocytes during in vitro culture
XX CC expansion, for preventing or treating bone-related or cartilage-related
XX CC diseases or Ets-related gene associated tumour, or as marker of
XX CC progenitor cells of stable chondrocytes. Cells expressing the
XX CC polynucleotide are useful for producing or repairing cartilage tissue
XX CC in a mammal.
XX SQ Sequence 1291 BP; 334 A; 409 C; 302 G; 246 T; 0 other;
XX Query Match 11.4%; Score 200.4; DB 25; Length 1291;
XX Best Local Similarity 77.4%; Pred. No. 2.7e-30;
XX Matches 243; Conservative 0; Mismatches 71; Indels 0; Gaps 0;
QY 529 AGCCCTGCGGTACAGAAAGGCGGAGCCGAGATCCAGTTGTGCGAGTTTCTACTGGAGCTG 588
Db 689 AGCCGCTTGCAATCCAGGAGTGCCAGATCCAGCTTTGCGAGTTCTCTCTGAGCTC 748
QY 589 CTGGCAGACCGCGCAACGCCGCTGTCATCGCTGGAGGGCGGCCAGCGGAGTTCAAG 648
Db 749 CTGTGGACAGCTCCAACTCCAGCTGCATCACCTGGGAAGGACCACCGGGAGTTCAAG 808
QY 649 CTCACGACCGCCGAGAGTGCGCGGACGCTGGGCGAGCGCAAGCAAGCCCAATATG 708
Db 809 ATGACGGATCCGACGAGGTGCCCGCGCTGGGAGAGCGGAGAGCAACCCAACTATG 868
QY 709 AACTACGACAAGCTAAGTCGACACTGCGCTACTACTACGACAAAACATCATGACGAAG 768
Db 869 AACTACGATAAGCTCAGCGCGCCCTCCGTTACTACTATGACAAACATCATGACCAAG 928
QY 769 GTGCAGCGCAGCGCTAGCGCTTACGCTTTGACTTCCAGGGCTCGGCAAGGCTTGCCAG 828
Db 929 GTCCATGGGAAGCGCTACGCTTACAAAGTTCGACTTCCAGCGGATCGCCCGAGCCCTCCAG 988

```

```

QY 829 CCACCACCGCGGCA 842
Db 989 CCCACCCCCCGGA 1002
RESULT 14
ABZ68769
ID ABZ68769 standard; DNA; 1372 BP.
XX AC ABZ68769;
XX DT 16-MAY-2003 (first entry)
XX DE Nucleotide sequence of human ERG splice variant C-1-4.
XX KW Human; splice variant; ERG; Ets; transcription factor; C-1-1; C-1-2;
XX KW C-1-3; C-1-4; chondrocyte; articular chondrocyte; bone disease;
XX KW cartilage disease; tumour; gene; ss.
XX OS Homo sapiens.
XX FH Key Location/Qualifiers
XX FT CDS 146..1309
XX FT FT /*tag= a
XX FT FT /product= "ERG splice variant C-1-4"
XX PN WO2003000724-A2.
XX PD 03-JAN-2003.
XX PF 08-MAR-2002; 2002WO-EP02605.
XX PR 25-JUN-2001; 2001US-300756P.
XX PA (TIGE-) TIGENIX NV.
XX PI Luyten F, De Bari C, Dell'Accio F;
XX DR WPI: 2003-221398/21.
XX DR P-PSDB; ABP97698.
XX PT Novel splice variants of the Ets-related gene (ERG) transcription
XX PT factor , useful for preventing or treating bone-related or
XX PT cartilage-related diseases or Ets-related gene associated tumour
XX PS Claim 112; Fig 7; 60pp; English.
XX CC The present sequence encodes splice variant C-1-4 of human ERG. ERG
XX CC belongs to the Ets family of transcription factors. The specification
XX CC describes C-1-1 (characterised by the absence of the 81 bp fragment, and
XX CC the presence of the 218 bp and 72 bp fragments of the ERG-3 cDNA
XX CC sequence); C-1-2 (characterised by the absence of the 81 bp and 218 bp
XX CC fragments, and the presence of the 72 bp fragment of the ERG-3 cDNA
XX CC sequence); C-1-3 (characterised by the absence of the 72 bp, 81 bp and
XX CC 218 bp fragments of the ERG-3 cDNA sequence); and C-1-4 (characterised by
XX CC the absence of the 218 bp fragment, and the presence of the 72 bp and
XX CC 81 bp fragments of the ERG-3 cDNA sequence). The splice variant
XX CC polynucleotides are useful for stabilizing the phenotype of chondrocytes
XX CC or cells with a chondrocytes-like phenotype, for preventing
XX CC de-differentiation of articular chondrocytes during in vitro culture
XX CC expansion, for preventing or treating bone-related or cartilage-related
XX CC diseases or Ets-related gene associated tumour, or as marker of
XX CC progenitor cells of stable chondrocytes. Cells expressing the
XX CC polynucleotide are useful for producing or repairing cartilage tissue
XX CC in a mammal.
XX SQ Sequence 1372 BP; 359 A; 430 C; 314 G; 269 T; 0 other;

```

```

Query Match 11.4%; Score 200.4; DB 25; Length 1372;
Best Local Similarity 77.4%; Pred. No. 2.7e-30;
Matches 243; Conservative 0; Mismatches 71; Indels 0; Gaps 0;

```

```
QY 529 AGCCCTCGGTACAGAAAGGCGGCGGAGATCCAGTTGTGGCAGTTTCTACTGAGCTG 588
    |||||
DB 770 AGCCGCTTGCAAAATCCAGGCGAGTGGCCAGATCCAGCTTTGGCAGTTTCTCTGGAGCTC 829
    |||||
QY 589 CTGGCAGACCGCGAAGCGCGGCTGCGTGGGAGGGCGGCCAGCGGAGTTCAAG 648
    |||||
DB 830 CTGTGGACAGCTCCAACTCCAGCTGCACTACCTGGGAAGGCCAACCGGGAGTTCAAG 889
    |||||
QY 649 CTCACCGACCGGACGAGGTGGCGGCGGAGCTGGGCGGAGCGCAAGCAAGCCCAATATG 708
    |||||
DB 890 ATGACGGATCCGACGAGGTGGCCGCGCTGGGCGAGAGCGCAAGCAAGCCCAACATG 949
    |||||
QY 709 AACTACGACAACTAAGTCGAGCAGTGGCTACTACTAGCAGCAAAACATCATGAGCAAG 768
    |||||
DB 950 AACTACGATAAGCTCAGCGCGCCCTCCGTTACTACTATGACAAAGAACATCATGACCAAG 1009
    |||||
QY 769 GTGACCGCAAGCGCTACGCTTACCGCTTTGACTTCCAGGGCTTGGCAGAGCTTGGCCAG 828
    |||||
DB 1010 GTCCATGGGAGCGGTAGCGCTACAGTTGCAAGTTGCACTTCCAGGGATGCCCGCCCTCCAG 1069
    |||||
QY 829 CCACACCGCGCA 842
    |||||
DB 1070 CCCCACCCCGGA 1083
    |||||
```

RESULT 15

ABX08759
ID ABX08759 standard; cDNA; 1389 BP.

AC ABX08759;

XX 21-JAN-2003 (first entry)

DE Angiogenesis-associated human polynucleotide sequence #21.

XX Human; angiogenesis-associated transcript; angiogenesis;

KW angiogenesis-associated disease; cancer; cytostatic; gene therapy;

KW gene; ss.

XX Homo sapiens.

OS
XX WO200279492-A2.

PN 10-OCT-2002.

XX 14-FEB-2002; 2002WO-US04915.

XX 14-FEB-2001; 2001US-0784356.

PR 22-FEB-2001; 2001US-0791390.

PR 19-APR-2001; 2001US-285475P.

PR 03-AUG-2001; 2001US-310025P.

PR 13-NOV-2001; 2001US-350666P.

PR 29-NOV-2001; 2001US-334244P.

XX (EOSB-) EOS BIOTECHNOLOGY INC.

XX Murray R, Glynn R, Watson SR, Aziz N;

XX WPI; 2003-040681/03.

XX P-PSDB; ABU03476.

XX Detecting angiogenesis-associated transcript in a cell for diagnosing
PT and treating cancer by contacting a sample with a polynucleotide that
PT exhibits changes in expression level as a function of time in tissue
PT undergoing angiogenesis -

PS Example 2; Page 200; 291pp; English.

XX The present invention relates to methods and compositions for

CC detecting an angiogenesis-associated transcript in a cell in

CC a patient. The method involves contacting a biological sample from

CC the patient with a polynucleotide that selectively hybridizes to a

CC sequence at least 80% identical to any of the angiogenesis-associated

CC human polynucleotide sequences given in the specification. These
CC angiogenesis-associated polynucleotide sequences comprise genes that
CC exhibit changes in expression levels as a function of time in tissue
CC undergoing angiogenesis. The method and the polynucleotide sequences
CC of the invention are useful for diagnosing and treating angiogenesis
CC and angiogenesis-associated diseases e.g. cancer. The polynucleotide
CC sequences are also useful in the gene therapy of such disorders. The
CC angiogenesis-associated proteins encoded by the polynucleotide
CC sequences are useful as a vaccine for therapeutic and prophylactic
CC immunisation. ABX08739-ABX08853 represent angiogenesis-associated
CC polynucleotide sequences.

XX Sequence 1389 BP; 359 A; 440 C; 323 G; 267 T; 0 other;

Query Match 11.4%; Score 200.4; DB 25; Length 1389;

Best Local Similarity 77.4%; Pred. No. 2.7e-30;

Matches 243; Conservative 0; Mismatches 71; Indels 0; Gaps 0;

QY 529 AGCCCTCGGTACAGAAAGGCGGCGGAGATCCAGTTGTGGCAGTTTCTACTGAGCTG 588

DB 850 AGCCGCTTGCAAAATCCAGGCGAGTGGCCAGATCCAGCTTTGGCAGTTTCTCTGGAGCTC 909

QY 589 CTGGCAGACCGCGCAAGCGCGGCTGCGTGGGAGGGCGGCCAGCGGAGTTCAAG 648

DB 910 CTGTGGACAGCTCCAACTCCAGCTGCACTACCTGGGAAGGCCAACCGGGAGTTCAAG 969

QY 649 CTCACCGACCGGCGGAGGTGGCGGCGGAGCTGGGCGGAGCGCAAGCAAGCCCAATATG 708

DB 970 ATGACGGATCCGACGAGGTGGCCGCGCTGGGAGAGCGGAGCAACCCCAACATG 1029

QY 709 AACTACGACAAAGCTAAGTCGAGCACTGGCTACTACTACGACAAACATCATGAGCAAG 768

DB 1030 AACTACGATAAGCTCAGCGCGCCCTCCGTTACTACTATGACAAAGACATCATGACCAAG 1089

QY 769 GTGACCGCAAGCGCTACGCTTACCGCTTTGACTTCCAGGGCTTGGCAGAGCTTGGCCAG 828

DB 1090 GTCCATGGGAAGCGGTAGCGCTACAGTTGCAAGTTGCACTTCCAGGGATGCCCGCCCTCCAG 1149

QY 829 CCACACCGCGCA 842

DB 1150 CCCCACCCCGGA 1163

Search completed: November 26, 2003, 13:33:03

Job time : 495 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 26, 2003, 13:24:47 ; Search time 571 Seconds
(without alignments)
10109.102 Million cell updates/sec

Title: US-10-027-859-1
Perfect score: 1752
Sequence: 1 ttgttttaacaaacattgtt.....aaaaaaaaaaaaaaaaaaaa 1752

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 2190069 seqs, 1647345023 residues

Total number of hits satisfying chosen parameters: 4380138

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications NA.*
1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/2/pubpna/PTCT_NEW_PUB.seq.*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq.*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq.*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq.*
6: /cgn2_6/ptodata/2/pubpna/PTCTUS_PUBCOMB.seq.*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq.*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq.*
9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq.*
10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq.*
11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq.*
12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq.*
13: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq.*
14: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq.*
15: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq.*
16: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq.*
17: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1752	100.0	1752	US-09-850-799-1	Sequence 1, Appli
2	1752	100.0	1752	US-10-027-859-1	Sequence 1, Appli
C 3	233.4	13.3	1754	US-10-091-333-1	Sequence 1, Appli
C 4	233.4	13.3	1754	US-10-325-878-1	Sequence 1, Appli
5	223	12.7	2532	US-10-037-270-869	Sequence 869, App
6	200.4	11.4	3166	US-10-021-660-26	Sequence 26, Appl
7	200.4	11.4	3166	US-10-025-823-98	Sequence 98, Appl
C 8	199.4	11.4	567	US-09-864-761-18410	Sequence 18410, A
C 9	198.4	11.3	473	US-09-864-761-20472	Sequence 20472, A
C 10	197.4	11.3	472	US-10-029-386-25224	Sequence 25224, A
11	194	11.1	2957	US-10-007-926A-295	Sequence 295, App
12	180	10.3	1447	US-09-902-772-1	Sequence 1, Appli
13	180	10.3	1528	US-09-902-772-3	Sequence 3, Appli
14	138.8	7.9	441	US-10-210-120-41	Sequence 41, Appl
15	138.8	7.9	473	US-09-918-995-27904	Sequence 27904, A
16	138.8	7.9	1884	US-09-925-300-420	Sequence 420, App

17	138.8	7.9	2188	10	US-09-920-300A-1716	Sequence 1716, Ap
18	138.8	7.9	2188	12	US-10-099-926-1716	Sequence 1716, Ap
19	138.8	7.9	2188	13	US-10-033-528-1716	Sequence 1716, Ap
20	138.8	7.9	2268	10	US-09-920-300A-1693	Sequence 1693, Ap
21	138.8	7.9	2268	12	US-10-099-926-1693	Sequence 1693, Ap
22	138.8	7.9	2268	13	US-10-033-528-1693	Sequence 1693, Ap
23	138.8	7.9	2269	10	US-09-954-531-955	Sequence 955, App
C 24	138.8	7.9	3692	14	US-10-106-698-824	Sequence 824, App
C 25	136.2	7.8	530	12	US-10-029-386-11517	Sequence 11517, A
26	132.2	7.5	3178	13	US-10-108-605-124	Sequence 124, App
27	132.2	7.5	3178	13	US-10-108-605-128	Sequence 128, App
28	125.4	7.2	549	9	US-09-923-779-54	Sequence 54, Appl
C 29	121	6.9	703	14	US-10-050-704-18	Sequence 18, Appl
C 30	121	6.9	1734	9	US-09-739-262-5	Sequence 5, Appli
C 31	121	6.9	1744	11	US-09-919-172-45	Sequence 45, Appl
C 32	121	6.9	1744	11	US-09-919-039-330	Sequence 330, App
C 33	121	6.9	1744	12	US-10-051-835-23	Sequence 23, Appl
C 34	121	6.9	1760	12	US-10-301-822-60	Sequence 60, Appl
C 35	121	6.9	1760	12	US-10-465-572-5	Sequence 5, Appli
C 36	121	6.9	1760	14	US-10-171-311-57	Sequence 57, Appl
C 37	121	6.9	1760	14	US-10-175-533-173	Sequence 173, Appl
C 38	121	6.9	1782	14	US-10-091-333-2	Sequence 2, Appli
C 39	121	6.9	1782	14	US-10-325-878-2	Sequence 2, Appli
C 40	119.4	6.8	1023	13	US-09-814-353-20061	Sequence 20061, A
C 41	118	6.7	366	10	US-10-002-600-2	Sequence 2, Appli
C 42	118	6.7	366	10	US-09-880-107-999	Sequence 999, App
C 43	118	6.7	366	12	US-09-873-319-188	Sequence 188, App
C 44	118	6.7	366	12	US-09-960-706-318	Sequence 318, App
C 45	118	6.7	366	12	US-09-873-367C-386	Sequence 386, App

ALIGNMENTS

RESULT 1

US-09-850-799-1
; Sequence 1, Application US/09850799
; Patent No. US20020090647A1
; GENERAL INFORMATION:
; APPLICANT: Deneris, Evan S.
; APPLICANT: Fyodorov, Dmitry V.
; APPLICANT: Hendricks, Timothy J.
; TITLE OF INVENTION: Reagents and Methods for the Screening of Compounds
; TITLE OF INVENTION: Useful in the Treatment of Neurological Diseases
; FILE REFERENCE: CASE-03828
; CURRENT APPLICATION NUMBER: US/09/850,799
; CURRENT FILING DATE: 2001-05-08 -
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/360,779
; PRIOR FILING DATE: EARLIER FILING DATE: 1999-07-26
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 1752
; TYPE: DNA
; ORGANISM: Rattus norvegicus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (112)..(1131)
US-09-850-799-1

Query Match	100.0%	Score 1752;	DB 9;	Length 1752;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 1752;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	TTTGTTTTAAACAACATGTTTATTAGAAAAGTAAATAATTCATAGGTCTTAGTACTTG	60	
DB	1	TTTGTTTTAAACAACATGTTTATTAGAAAAGTAAATAATTCATAGGTCTTAGTACTTG	60	
QY	61	AACATCAAGTGTATTTCATGAACCGTGACTTCTTTCATCTTAACACGTTCTTAGATGAAGAC	120	
DB	61	AACATCAAGTGTATTTCATGAACCGTGACTTCTTTCATCTTAACACGTTCTTAGATGAAGAC	120	
QY	121	CCAGTGGCGTCTCTCTGGGGGAGAGGGTTCACAGCCCCCCCCCTCCAGCCCCATCCC	180	

Db 121 CCAGGTGGCGCTCTCTGGGGAGAGGTTCCAGCCCCCAGCCGCCCTCAGCCCCATCCC 180
Qy 181 CTCACAGCTCACTCTCTCCAGTACACCGGCACCGGATGGCTGGGATGAGTCCAGGAC 240
Db 181 CTCACAGCTCACTCTCTCCAGTACACCGGCACCGGATGGCTGGGATGAGTCCAGGAC 240
Qy 241 CCGCTCCCTCTTCCACACACCTTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGG 300
Db 241 CCGCTCCCTCTTCCACACACCTTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGG 300
Qy 301 GGGTCTACTCTTCCCTGTCCACACAGTCCACACAGTCCACACAGTCCACACAGTCC 360
Db 301 GGGTCTACTCTTCCCTGTCCACACAGTCCACACAGTCCACACAGTCCACACAGTCC 360
Qy 361 AAGCTGACGTGCCCCCGCTGGGTCCTCCCGCCATCTCCACCGCCAGTCCCGGAGCG 420
Db 361 AAGCTGACGTGCCCCCGCTGGGTCCTCCCGCCATCTCCCGCCAGTCCCGGAGCG 420
Qy 421 ATGAGACAGAGCGGACCTCCAGCGCCCTGCTGATCAACATGATCTTACAGATCCGTC 480
Db 421 ATGAGACAGAGCGGACCTCCAGCGCCCTGCTGATCAACATGATCTTACAGATCCGTC 480
Qy 481 GGAGATGGTCTTTTAAAGAAAGGAGAGCCGAGCTGGGGCCGCTGAGCCCTCGGTA 540
Db 481 GGAGATGGTCTTTTAAAGAAAGGAGAGCCGAGCTGGGGCCGCTGAGCCCTCGGTA 540
Qy 541 CAGAAAGGAGCGGGGAGATCCAGTTGTGGCAGTTTCTACTGGAGCTGTGGCAGACGC 600
Db 541 CAGAAAGGAGCGGGGAGATCCAGTTGTGGCAGTTTCTACTGGAGCTGTGGCAGACGC 600
Qy 601 GCGACCGCGCTGATCCGCTGGAGGGGCGCCAGCGGAGTTCAAGTCTACCAACCCC 660
Db 601 GCGACCGCGCTGATCCGCTGGAGGGGCGCCAGCGGAGTTCAAGTCTACCAACCCC 660
Qy 661 GACGAGTGGCGGACCGCTGGGGCGGAGCGCAAGCAAGCCCAATATGAATACGACAAG 720
Db 661 GACGAGTGGCGGACCGCTGGGGCGGAGCGCAAGCAAGCCCAATATGAATACGACAAG 720
Qy 721 CTAAGTCTGAGCACTGGCTTACTACTACGACAAACATCATGAGCAAGTGCACGCAAG 780
Db 721 CTAAGTCTGAGCACTGGCTTACTACTACGACAAACATCATGAGCAAGTGCACGCAAG 780
Qy 781 CGCTACGCTACCGCTTTGACTTCCAGGGCTGGCAGAGCTTGCAGCCACACCCCGG 840
Db 781 CGCTACGCTACCGCTTTGACTTCCAGGGCTGGCAGAGCTTGCAGCCACACCCCGG 840
Qy 841 CACGCCACCGCGCTGCGCGCGCGCGGAGCGCGCGCGCGCGCGCGCGCGCGCGCTT 900
Db 841 CACGCCACCGCGCTGCGCGCGCGCGGAGCGCGCGCGCGCGCGCGCGCGCGCGCTT 900
Qy 901 TACAAAGCTCCGGCTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 960
Db 901 TACAAAGCTCCGGCTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 960
Qy 961 ATGGAGCGCTCGCGCGCGCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1020
Db 961 ATGGAGCGCTCGCGCGCGCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1020
Qy 1021 ACCCGCGCTGCGCGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1080
Db 1021 ACCCGCGCTGCGCGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1080
Qy 1081 GGGCGCTTTGGCGCGCTGCGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCG 1140
Db 1081 GGGCGCTTTGGCGCGCTGCGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCG 1140
Qy 1141 CCGCGCGCTGCGCGCGCTGCGCGCTGCGCGCTGCGCGCTGCGCGCTGCGCGCTGCGCG 1200
Db 1141 CCGCGCGCTGCGCGCGCTGCGCGCTGCGCGCTGCGCGCTGCGCGCTGCGCGCTGCG 1200
Qy 1201 GAGGAGCGCGGAGATTTCCCGAGCGTTCTTTTACACAGATTTTCGTTGCGAGCGCGCT 1260

Db 1201 GAGGAGCGCGGAGATTTCCCGAGCGTTCTTTTACACAGATTTTCGTTGCGAGCGCGCT 1260
Qy 1261 CCCAGCCAGGGAAGAAAGGATGGGAAGCCTCTGAGGCTCTTCTTGAATACGAGGCTTCC 1320
Db 1261 CCCAGCCAGGGAAGAAAGGATGGGAAGCCTCTGAGGCTCTTCTTGAATACGAGGCTTCC 1320
Qy 1321 AGGCTCCCATATCATCATCAACAGGAGGTCATGTCCTCCACCTTTAAATTTTCTCTT 1380
Db 1321 AGGCTCCCATATCATCATCAACAGGAGGTCATGTCCTCCACCTTTAAATTTTCTCTT 1380
Qy 1381 CCAAGTCTCCAGATTTCTGMACTCCCGTCTTTTCTTCTTCTTCTTCTTCTTCTTCTT 1440
Db 1381 CCAAGTCTCCAGATTTCTGMACTCCCGTCTTTTCTTCTTCTTCTTCTTCTTCTTCTT 1440
Qy 1441 CTTCTCTTTATGACCCCTAGTTTCTGTTTGTGTTTTTTTTTTTTTTTTTTTTTCTCTC 1500
Db 1441 CTTCTCTTTATGACCCCTAGTTTCTGTTTGTGTTTTTTTTTTTTTTTTTTTTTCTCTC 1500
Qy 1501 ATTTTCTTCTTCCAGACCTACTCCAAAGGATAGTACCTCGGTAGTACCTCGAGGCT 1560
Db 1501 ATTTTCTTCTTCCAGACCTACTCCAAAGGATAGTACCTCGGTAGTACCTCGAGGCT 1560
Qy 1561 TCTCACACTCCCTTTTCCGGATATGAGAAGCATCAAAAACATCTCTGCTGTTGTCATC 1620
Db 1561 TCTCACACTCCCTTTTCCGGATATGAGAAGCATCAAAAACATCTCTGCTGTTGTCATC 1620
Qy 1621 CCTATCCCAACACTCTGCTTTCCTTCCATACCACTCTTCCGCGCAAGGACCTCG 1680
Db 1621 CCTATCCCAACACTCTGCTTTCCTTCCATACCACTCTTCCGCGCAAGGACCTCG 1680
Qy 1681 TCTGTATATATTCCTTTTCCAGCCCATTAAGATCCAAAGGATCCAAAGGATCCAAAG 1740
Db 1681 TCTGTATATATTCCTTTTCCAGCCCATTAAGATCCAAAGGATCCAAAGGATCCAAAG 1740
Qy 1741 AAAAAAAAAA 1752
Db 1741 AAAAAAAAAA 1752

RESULT 2
US-10-027-859-1
; Sequence 1, Application US/10027859
; Publication No. US20030175930A1
; GENERAL INFORMATION:
; APPLICANT: Deneris, Evan S.
; APPLICANT: Hendricks, Timothy J.
; TITLE OF INVENTION: Reagents and Methods for the Screening of Compounds
; TITLE OF INVENTION: Useful in the Treatment of Neurological Diseases
; FILE REFERENCE: CASE-04027
; CURRENT APPLICATION NUMBER: US/10/027,859
; CURRENT FILING DATE: 2001-10-25
; PRIOR APPLICATION NUMBER: US/09/435,335
; PRIOR FILING DATE: 1999-11-05
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/360,779
; PRIOR FILING DATE: EARLIER FILING DATE: 1999-07-26
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 1752
; TYPE: DNA
; ORGANISM: Rattus norvegicus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (112) .. (1131)
US-10-027-859-1

Query Match 100.0%; Score 1752; DB 12; Length 1752;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1752; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TTTGTTTAAACAAACATGTTTATTAGAAAAGTAAATAATTGATAGGCTTAGTACTTG 60

Db 1 TTTGTTTAAACAATGTTTATTAGAAAAAGTAAATAATTGCATAGGTCTTAGTACTTG 60
Qy 61 AACATCAAGTGATTTATGAACCGTGAATATCTTTCATGTAACAGTCTTAGATGGAAGAC 120
Db 61 AACATCAAGTGATTTATGAACCGTGAATATCTTTCATGTAACAGTCTTAGATGGAAGAC 120
Qy 121 CCAGGTGGCGTCTCTCTGGGGAGAGGTTTCCAGCCGCCCAACCCCTCAGCCGCAATCCC 180
Db 121 CCAGGTGGCGTCTCTCTGGGGAGAGGTTTCCAGCCGCCCAACCCCTCAGCCGCAATCCC 180
Qy 181 CTCACAGTCACTCTCTCCAGTACACCGGACCGGGATGGGTGGGATCGAGCTCCAGGAC 240
Db 181 CTCACAGTCACTCTCTCCAGTACACCGGACCGGGATGGGTGGGATCGAGCTCCAGGAC 240
Qy 241 CCCCTCCCTCTCACCACACCCCTGCTGCCCGCTCCCGCCAGGCTTGC CGGACCCGCG 300
Db 241 CCCCTCCCTCTCACCACACCCCTGCTGCCCGCTCCCGCCAGGCTTGC CGGACCCGCG 300
Qy 301 CGGTCTACTCTCTCTGTCACCCACAGTACACCGGGGGGTATCGGCACCCCAAGCGCA 360
Db 301 CGGTCTACTCTCTCTGTCACCCACAGTACACCGGGGGGTATCGGCACCCCAAGCGCA 360
Qy 361 AAGCTGACGTGCCCGCCCGTGGTCCCGCCATCTCCACCGCCAGTCCCGGACGCG 420
Db 361 AAGCTGACGTGCCCGCCCGTGGTCCCGCCATCTCCACCGCCAGTCCCGGACGCG 420
Qy 421 ATGACAGAGCGGCACTCCAGCCCTGCTGATCAACATGTATCTACACAGATCCCGTC 480
Db 421 ATGACAGAGCGGCACTCCAGCCCTGCTGATCAACATGTATCTACACAGATCCCGTC 480
Qy 481 GGAGATGGTCTTTTAAAGGAAGGAGAGCCGAGCTGGGGCGCTGAGCCCTGCGGTA 540
Db 481 GGAGATGGTCTTTTAAAGGAAGGAGAGCCGAGCTGGGGCGCTGAGCCCTGCGGTA 540
Qy 541 CAGAAAGCGACGGGAGATCAGTTGTCAGATTTCTACTGGAGCTGCTGGCAGACCGC 600
Db 541 CAGAAAGCGACGGGAGATCAGTTGTCAGATTTCTACTGGAGCTGCTGGCAGACCGC 600
Qy 601 GCGAACCGCGCTGCATCGCTGGAGGGCGGCCACCGCGAGTTCAAGCTCACCGACCCC 660
Db 601 GCGAACCGCGCTGCATCGCTGGAGGGCGGCCACCGCGAGTTCAAGCTCACCGACCCC 660
Qy 661 GACGAGTGGCGGACGCTGGGGAGCGGCAAGAGCAAGCCCAATATGAACATGACAAAG 720
Db 661 GACGAGTGGCGGACGCTGGGGAGCGGCAAGAGCAAGCCCAATATGAACATGACAAAG 720
Qy 721 CTAAGTCAGCACTGCGCTACTACTACGACAAAACATCATGACAGGTCGACGCGCAAG 780
Db 721 CTAAGTCAGCACTGCGCTACTACTACGACAAAACATCATGACAGGTCGACGCGCAAG 780
Qy 781 CGCTACGCTTACCGCTTTGACTTCCAGGGCTGGCAGAGGTTGCCAGCCACCAACCCGCG 840
Db 781 CGCTACGCTTACCGCTTTGACTTCCAGGGCTGGCAGAGGTTGCCAGCCACCAACCCGCG 840
Qy 841 CACGCCACGCGCGCTGCGCGCCGCGCAGCGCGCGCGCGCGCGCGCGCGCGCTT 900
Db 841 CACGCCACGCGCGCTGCGCGCCGCGCAGCGCGCGCGCGCGCGCGCGCGCGCTT 900
Qy 901 TACAAGTCCCGGCTGGTCTGGCTCCACTGCGCTTCCCGGCGCTTCCAAACTCAACCTT 960
Db 901 TACAAGTCCCGGCTGGTCTGGCTCCACTGCGCTTCCCGGCGCTTCCAAACTCAACCTT 960
Qy 961 ATGCAGCTCGCGCGGCTGGCGCGCTGCTGCTTCTTACTGGCGCTGGTCCCAAGCGC 1020
Db 961 ATGCAGCTCGCGCGGCTGGCGCGCTGCTGCTTCTTACTGGCGCTGGTCCCAAGCGC 1020
Qy 1021 ACCGCGCTCGCGCGGCTGCGCGCTGCTGCTTACCCAAACCGCGCTGCGCGCTTCC 1080
Db 1021 ACCGCGCTCGCGCGGCTGCGCGCTGCTGCTTACCCAAACCGCGCTGCGCGCTTCC 1080
Qy 1081 GGGCCCTTTGGCGGCTGGCGCGCTTGGCACTTGGGGGTCAATTACTAGACGGGA 1140
Db 1081 GGGCCCTTTGGCGGCTGGCGCGCTTGGCACTTGGGGGTCAATTACTAGACGGGA 1140

Qy 1141 CGGCCGGTGCAGTGGGGCTCTCCACACAGCCAGTGACCAATCCCATCTCATCTGG 1200
Db 1141 CGGCCGGTGCAGTGGGGCTCTCCACACAGCCAGTGACCAATCCCATCTCATCTGG 1200
Qy 1201 GAGGAGCCCCGAAGATTTCCCGAGGTTCTTTTACCACAGATTTTGGTGGCAGCCGCT 1260
Db 1201 GAGGAGCCCCGAAGATTTCCCGAGGTTCTTTTACCACAGATTTTGGTGGCAGCCGCT 1260
Qy 1261 CCCAGCCAGGGAAGAGATGGAAGCCCTCAGGCTTCTTGAATACGAGGCTTCC 1320
Db 1261 CCCAGCCAGGGAAGAGATGGAAGCCCTCAGGCTTCTTGAATACGAGGCTTCC 1320
Qy 1321 AGGCTCCCATTTATCATCACCCAGGAGGTCATGTGCTCCCACTTTTAACTTTCTCTT 1380
Db 1321 AGGCTCCCATTTATCATCACCCAGGAGGTCATGTGCTCCCACTTTTAACTTTCTCTT 1380
Qy 1381 CCAAGTCTCAGATTTCTGGAATCCCGTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCT 1440
Db 1381 CCAAGTCTCAGATTTCTGGAATCCCGTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCT 1440
Qy 1441 CTTCCTCTTTATGACCCCTAGTTTCTGTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCT 1500
Db 1441 CTTCCTCTTTATGACCCCTAGTTTCTGTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCT 1500
Qy 1501 ATTTTCTTCTCTCCACGACCTACTCCAAACGGTAGTACCTCGGTAGTACCTCGAGGCT 1560
Db 1501 ATTTTCTTCTCTCCACGACCTACTCCAAACGGTAGTACCTCGGTAGTACCTCGAGGCT 1560
Qy 1561 TCTCACACTCTCCCTTTTTCGGGATATGAGAGCATCAAAAAATCTCTCTGTTGTCATC 1620
Db 1561 TCTCACACTCTCCCTTTTTCGGGATATGAGAGCATCAAAAAATCTCTCTGTTGTCATC 1620
Qy 1621 CCTATCCCAACACTCTGCTCGCTCCCTTCCATACACACTCTTGCCCAAGGACCTCG 1680
Db 1621 CCTATCCCAACACTCTGCTCGCTCCCTTCCATACACACTCTTGCCCAAGGACCTCG 1680
Qy 1681 TCTGTATATATCTTTTTCAGCCCCATTAAGATCAAGCTTCAAAAAAAGGACCTTCA 1740
Db 1681 TCTGTATATATCTTTTTCAGCCCCATTAAGATCAAGCTTCAAAAAAAGGACCTTCA 1740
Qy 1741 AAAAAAAGGACCTTCAAGCTTCAAGCTTCAAGCTTCAAGCTTCAAGCTTCAAGCTTCA 1740
Db 1741 AAAAAAAGGACCTTCAAGCTTCAAGCTTCAAGCTTCAAGCTTCAAGCTTCAAGCTTCA 1740

RESULT 3
US-10-091-333-1/c
; Sequence 1, Application US/10091333
; Publication No. US20030104973A1
; GENERAL INFORMATION:
; APPLICANT: EINAT, Paz
; APPLICANT: SKALITER, Rami
; TITLE OF INVENTION: HYPOXIA-REGULATED GENES
; FILE REFERENCE: EINAT-1.1D
; CURRENT APPLICATION NUMBER: US/10/091.333
; CURRENT FILING DATE: 2002-03-06
; PRIOR APPLICATION NUMBER: US 09/604,978
; PRIOR FILING DATE: 2000-06-28
; PRIOR APPLICATION NUMBER: US 09/138,112
; PRIOR FILING DATE: 1999-08-21
; PRIOR APPLICATION NUMBER: US 60/056,453
; PRIOR FILING DATE: 1997-08-21
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 1754
; TYPE: DNA
; ORGANISM: Rattus norvegicus
US-10-091-333-1

Query Match 13.3%; Score 233.4; DB 14; Length 1754;
Best Local Similarity 99.6%; Pred. No. 2.8e-51;

Matches 234; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TTGTTTTTAAACAACATGTTTATTAGAAAAGTAAAAATATTGTCATAGGCTTTAGTACTTG 60
Db 1733 TTGTTTTTAAACAACATGTTTATTAGAAAAGTAAAAATATTGTCATAGGCTTTAGTACTTG 1674
QY 61 AACATCAAGTGATTTCATGAACCGTGAGTATCTTTCATGTAAACAGTTCTAGATGGAAGAC 120
Db 1673 AACATCAAGTGATTTCATGAACCGTGAGTATCTTTCATGTAAACAGTTCTAGATGGAAGAC 1614
QY 121 CCAGGTGGCGCTCTCTGGGGGAGAGGGTTCCAGCCCCCAGCCCCCTCAGCCCCATCCC 180
Db 1613 CCAGGTGGCGCTCTCTGGGGGAGAGGGTTCCAGCCCCCAGCCCCCTCAGCCCCATCCC 1554
QY 181 CTCACAGCTCACTCTCCAGTACACCGGCACCGGATGGCTGGGATGCAGCTCC 235
Db 1553 CTCACAGTTCACTCTCCAGTACACCGGCACCGGATGGCTGGGATGCAGCTCC 1499

RESULT 4

US-10-325-878-1/c
; Sequence 1, Application US/10325878
; Publication No. US20030124116A1
; GENERAL INFORMATION:
; APPLICANT: Einat, Paz
; TITLE OF INVENTION: HYPKXIA-REGULATED GENES
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: KOHN & ASSOCIATES
; STREET: 30500 No. US20030124116A1althwestern Hwy., Suite 401
; CITY: Farmington Hills
; STATE: Michigan
; COUNTRY: U.S.
; ZIP: 48334
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/325,878
; FILING DATE: 23-Dec-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA: US/09/138,112
; ATTORNEY/AGENT INFORMATION:
; NAME: Kohn, Kenneth I.
; REGISTRATION NUMBER: 30,955
; REFERENCE/DOCKET NUMBER: 0168.00034
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (248) 539-5050
; TELEFAX: (248) 5395055
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1754 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-10-325-878-1

Query Match 13.3%; Score 233.4; DB 14; Length 1754;
Best Local Similarity 99.6%; Pred. No. 2.8e-51;
Matches 234; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 TTGTTTTTAAACAACATGTTTATTAGAAAAGTAAAAATATTGTCATAGGCTTTAGTACTTG 60
Db 1733 TTGTTTTTAAACAACATGTTTATTAGAAAAGTAAAAATATTGTCATAGGCTTTAGTACTTG 1674

QY 61 AACATCAAGTGATTTCATGAACCGTGAGTATCTTTCATGTAAACAGTTCTAGATGGAAGAC 120
Db 1673 AACATCAAGTGATTTCATGAACCGTGAGTATCTTTCATGTAAACAGTTCTAGATGGAAGAC 1614
QY 121 CCAGGTGGCGCTCTCTGGGGGAGAGGGTTCCAGCCCCCAGCCCCCTCAGCCCCATCCC 180
Db 1613 CCAGGTGGCGCTCTCTGGGGGAGAGGGTTCCAGCCCCCAGCCCCCTCAGCCCCATCCC 1554
QY 181 CTCACAGCTCACTCTCCAGTACACCGGCACCGGATGGCTGGGATGCAGCTCC 235
Db 1553 CTCACAGTTCACTCTCCAGTACACCGGCACCGGATGGCTGGGATGCAGCTCC 1499

RESULT 5

US-10-037-270-869
; Sequence 869, Application US/10037270
; Publication No. US20030104529A1
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Zhang, Jie
; APPLICANT: Ren, Feiyan
; APPLICANT: Chen, Rui-hong
; APPLICANT: Zhao, Qing A.
; APPLICANT: Wehrman, Tom
; APPLICANT: Xue, Aidong J.
; APPLICANT: Yang, Yonghong
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Zhou, Ping
; APPLICANT: Ma, Yunqing
; APPLICANT: Wang, Dunrui
; APPLICANT: Wang, Zhiwei
; APPLICANT: Tillinghast, John
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: No. US20030104529A1el Nucleic Acids and
; FILE OF INVENTION: Polypeptides
; FILE REFERENCE: 784CIP28
; CURRENT APPLICATION NUMBER: US/10/037,270
; PRIOR FILING DATE: 2002-01-04
; PRIOR APPLICATION NUMBER: 09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/488,725
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 1104
; SOFTWARE: pt_FL_genes Version 1.0
; SEQ ID NO 869-
; LENGTH: 2532
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(2532)
US-10-037-270-869

Query Match 12.7%; Score 223; DB 14; Length 2532;
Best Local Similarity 72.2%; Pred. No. 1.9e-48;
Matches 366; Conservative 0; Mismatches 70; Indels 71; Gaps 3;
QY 546 AGGCAGCGGCAGATCCAGTTGTGGCAGTTTCTACTGGAGCTGTGGCAGACCGCGCGAA 605
Db 2096 ATGCAGCGGACAGATCCAGCTGTGGCAGTTTCTGTGGAGCTGTGGCTGACCGCGCGAA 2155

QY 605 CGCCGGCTGTCATCGCTGGGAGCGGCACCGCAGTTCAAGCTCACCGACCCCGACCA 665
Db 2156 CG-CGGCTGTCATCGCTGGGAGCGGTACCGCGAGTTCAAGCTCACCGACCGCA 2214
QY 666 GTTGGCGGCACCTCGTGGGCGAGCGCAAGCAAGCCCAATATGAACATACGAAGTAAG 725
Db 2215 GGTGGCGCGCGGTGGGCGAGCGCAAGCAAGCAAGCAACATGAACATACGAAGTGAG 2274
QY 726 TCGAGCACTCGCCTACTACTACGACAAAAACATCATGACAGGTGCGACGGCAAGCGCTA 785

Db 2275 CGCGCCCTGCTAT-CTACTGACAAAGAACATCATGACAGGTCATGTAAGCGCTA 2333
QY 786 CGCCTACCGCTTTGACTTCCAGGCGCTGGCACAGGTTGCCAGCCACACCGCGCACGC 845
Db 2334 CGCCTACCGCTTCGACTTCCAGGCGCTG-----2361
QY 846 CCACGCGCGCGCTGCGCGCGCGCAGCGGACGCGCGCCCGCCAGGATGGCGCACTTTACAA 905
Db 2362 -----GGCAGGACGCGCGCTCTACAA 2384
QY 906 GCTCCGCGCTGCTGGCTCCACTGCTTCCCGGCTCTCCAACTCAACCTTATGCG 965
Db 2385 GCTGCCCGCGCGCTCGACCGCTGCTTCCCGGCTCTCCAACTCAACCTATGCG 2444
QY 966 AGCCTCGCGCGCGCTGGCGCGCGCTGCTTCTTACTGGCTGTGCCAAGCCACCGC 1025
Db 2445 CGCCTCGCTGGAGTTCGCGCGCGCTTCTCTACTGCGCGCGGTGCGCGCCAC 2504
QY 1026 CGCTCGCGCGCGCACCGCTGCGCTCTA 1052
Db 2505 CGCTGAAGCGCGCAGCGCGCGCTCTA 2531

RESULT 6

US-10-021-660-26
; Sequence 26, Application US/10021660
; Publication No. US20030152926A1
; GENERAL INFORMATION:
; APPLICANT: Murray, Richard
; APPLICANT: Glynné, Richard
; APPLICANT: Watson, Susan R.
; APPLICANT: EOS Biotechnology, Inc.
; TITLE OF INVENTION: No. US20030152926A1 Methods of Diagnosis of Angiogenesis,
; TITLE OF INVENTION: Compositions and Methods of Screening for Angiogenesis
; TITLE OF INVENTION: Modulators
; FILE REFERENCE: 018501-00071005
; CURRENT APPLICATION NUMBER: US/10/021,660
; PRIOR FILING DATE: 2001-12-06
; PRIOR APPLICATION NUMBER: US/09/784,356
; PRIOR FILING DATE: 2001-02-14
; PRIOR APPLICATION NUMBER: US 09/637,977
; PRIOR FILING DATE: 2000-08-11
; NUMBER OF SEQ ID NOS: 135
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 26
; LENGTH: 3166
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-021-660-26

Query Match 11.4%; Score 200.4; DB 12; Length 3166;
Best Local Similarity 77.4%; Pred. No. 2.1e-42;
Matches 243; Conservative 0; Mismatches 71; Indels 0; Gaps 0;
QY 529 AGCCCTGGGTACAGAAAGGCGGCGAGATCCAGTTGTGGCACTTTCTACTGGAGCTG 588
Db 1106 AGCCGCTTGCAGAAATCCAGGAGTGGCCAGATCCAGCTTTGGCAGTTCTCTCGAGCTC 1165
QY 589 CTGGCAGACCGCGCAACCGCGCTGTCATCGCTGGGAGGCGGCCACCGCGAGTTCAAG 648
Db 1166 CTGTGGACAGCTCCAACTCCAGCTGTCATCACCTGGGAAGGCACCAACGGGAGTTCAAG 1225
QY 649 CTCACCGACCGCGAGGTGGCGGAGCGCTGGGCGAGCGCAAGCAAGCCCAATATG 708
Db 1226 ATGACGGATCCGACGAGGTGGCGCGCTGGGAGAGCGGAGCAACACCCCAATG 1285
QY 709 AACTACGACAGCTTAAGTCGAGCATCGCTACTACTAGCAAAACATCATGAGCAAG 768
Db 1286 AACTACGATAAGCTCAGCGCGCTCGGTTACTACTATGACAAACATCATGACCAAG 1345
QY 769 GTGCACGCAAGCGCTACGCTTACCGCTTTGACTTCCAGGCGCTGGCACAGGCTTGCAG 828

Db 1346 GTCCATGGGAAGCGCTACGCTTACAAAGTTCCAGGATCGCCACGCGCTCCAG 1405
QY 829 CCACCACCGCGCA 842
Db 1406 CCCACCCCCCGA 1419
RESULT 7
US-10-205-823-98
; Sequence 98, Application US/10205823
; Publication No. US20030108963A1
; GENERAL INFORMATION:
; APPLICANT: Schlegel, Robert
; APPLICANT: Monahan, John E.
; APPLICANT: Endege, Wilson O.
; APPLICANT: Gannavarapu, Manjula
; APPLICANT: Gorbacheva, Bella
; APPLICANT: Hoerssch, Sebastian
; APPLICANT: Kamatkar, Shubhangi
; APPLICANT: Wonsay, Angela M.
; APPLICANT: Glatt, Karen
; APPLICANT: Zhao, Xumei
; APPLICANT: Anderson, Dustin
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND
; TITLE OF INVENTION: METHODS FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF PROSTATE CANCER
; FILE REFERENCE: MRI-044
; CURRENT APPLICATION NUMBER: US/10/205,823
; CURRENT FILING DATE: 2002-07-25
; PRIOR APPLICATION NUMBER: 60/307,982
; PRIOR FILING DATE: 2001-07-25
; PRIOR APPLICATION NUMBER: 60/314,356
; PRIOR FILING DATE: 2001-08-22
; PRIOR APPLICATION NUMBER: 60/325,020
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: 60/341,746
; PRIOR FILING DATE: 2001-12-12
; PRIOR APPLICATION NUMBER: 60/362,158
; PRIOR FILING DATE: 2002-03-05
; NUMBER OF SEQ ID NOS: 455
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 98
; LENGTH: 3166
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-205-823-98

Query Match 11.4%; Score 200.4; DB 14; Length 3166;
Best Local Similarity 77.4%; Pred. No. 2.1e-42;
Matches 243; Conservative 0; Mismatches 71; Indels 0; Gaps 0;
QY 529 AGCCCTGGGTACAGAAAGGCGGCGAGATCCAGTTGTGGCACTTTCTACTGGAGCTG 588
Db 1106 AGCCGCTTGCAGAAATCCAGGAGTGGCCAGATCCAGCTTTGGCAGTTCTCTCGAGCTC 1165
QY 589 CTGGCAGACCGCGCAACCGCGCTGTCATCGCTGGGAGGCGGCCACCGCGAGTTCAAG 648
Db 1166 CTGTGGACAGCTCCAACTCCAGCTGTCATCACCTGGGAAGGCACCAACGGGAGTTCAAG 1225
QY 649 CTCACCGACCGCGAGGTGGCGGAGCGCTGGGCGAGCGCAAGCAAGCCCAATATG 708
Db 1226 ATGACGGATCCGACGAGGTGGCGCGCTGGGAGAGCGGAGCAACACCCCAATG 1285
QY 709 AACTACGACAGCTTAAGTCGAGCATCGCTACTACTAGCAAAACATCATGAGCAAG 768
Db 1286 AACTACGATAAGCTCAGCGCGCTCGGTTACTACTATGACAAACATCATGACCAAG 1345
QY 769 GTGCACGCAAGCGCTACGCTTACCGCTTTGACTTCCAGGCGCTGGCACAGGCTTGCAG 828
Db 1346 GTCCATGGGAAGCGCTACGCTTACCAAGTTCCAGGCGCTGGCACAGGCTTGCAG 1405
QY 829 CCACCACCGCGCA 842

Db 1406 CCCCCCCCCGGA 1419

RESULT 8

US-09-864-761-18410/c

; Sequence 18410, Application US/09864761

; Patent No. US20020048763A1

GENERAL INFORMATION:

; APPLICANT: Penn, Sharron G.

; APPLICANT: Rank, David R.

; APPLICANT: Hanzel, David K.

; APPLICANT: Chen, Wensheng

; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR

; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY

; FILE REFERENCE: Aecmica-x-1

; CURRENT APPLICATION NUMBER: US/09/864,761

; CURRENT FILING DATE: 2001-05-23

; PRIOR APPLICATION NUMBER: US 60/180,312

; PRIOR FILING DATE: 2000-02-04

; PRIOR APPLICATION NUMBER: US 60/207,456

; PRIOR FILING DATE: 2000-05-26

; PRIOR APPLICATION NUMBER: US 09/632,366

; PRIOR FILING DATE: 2000-08-03

; PRIOR APPLICATION NUMBER: GB 24263.6

; PRIOR FILING DATE: 2000-10-04

; PRIOR APPLICATION NUMBER: US 60/236,359

; PRIOR FILING DATE: 2000-09-27

; PRIOR APPLICATION NUMBER: PCT/US01/00666

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: PCT/US01/00667

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: PCT/US01/00664

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: PCT/US01/00669

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: PCT/US01/00665

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: PCT/US01/00668

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: PCT/US01/00663

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: PCT/US01/00662

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: PCT/US01/00661

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: PCT/US01/00670

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: US 60/234,687

; PRIOR FILING DATE: 2000-09-21

; PRIOR APPLICATION NUMBER: US 09/608,408

; PRIOR FILING DATE: 2000-06-30

; PRIOR APPLICATION NUMBER: US 09/774,203

; PRIOR FILING DATE: 2001-01-29

NUMBER OF SEQ ID NOS: 49117

SOFTWARE: Annonmax Sequence Listing Engine vers. 1.1

SEQ ID NO 18410

LENGTH: 567

TYPE: DNA

ORGANISM: Homo sapiens

FEATURE:

OTHER INFORMATION: MAP TO AP000163.1

OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 2.4

OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 1.4

OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.7

OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 0.92

OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 2.1

OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.8

OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.5

OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.5

OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 1

OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 0.95

OTHER INFORMATION: SWISSPLOT HIT: P11308, EVALUE 9.00e-99

OTHER INFORMATION: EST_HUMAN HIT: R87572.1, EVALUE 0.00e+00

; OTHER INFORMATION: NT HIT: M17254.1, EVALUE 0.00e+00

US-09-864-761-18410

Query Match 11.4%; Score 199.4; DB 9; Length 567;

Best Local Similarity 79.5%; Pred. No. 1.7e-42;

Matches 236; Conservative 0; Mismatches 61; Indels 0; Gaps 0;

QY 546 AGGCAGCGGCAGATCCAGTTGTGGCAGTTTCTACTGGAGCTGCTGGCAGACCGCGGAA 605

DB 523 AGGCAGTGGCCAGATCCAGCTTTGGCAGTTTCTCTCTGGAGCTCTCTTCGCGACAGCTCAA 464

QY 606 CCGCCGCTCATCGCTGGAGGGGGCCACCGCGAGTTCAAGCTCACCAGACCCCGACGA 665

DB 463 CTCACGCTCATCACCTGGAAAGGCCAACCGGGAGTTCAAGATACGGATCCCCGACGA 404

QY 666 GGTGGCGGACGCTGGGGCGAGCGCAAGAGCAAGCCCAATATGAACTACGACAAGCTAAG 725

DB 403 GGTGGCCCGGCGCTGGGGAGAGCGGAAGACAAACCAACATGAACTACGATAAGCTCAG 344

QY 726 TCGAGCACTGCGCTACTACTACGACAAAACATCATGACGACAGGTGCACGGCAAGCGCTA 785

DB 343 CCGCGCCCTCCGTTACTACTATGACAAAGAACATCATGACCAAGGTCCATGGGAAGCGCTA 284

QY 786 CCGCTACCGCTTTGACTTCCAGGGCTGGCAGAGGCTTCCAGCAGCACCCCGCGCA 842

DB 283 CCGCTACAAAGTTTCGACTTCCACGGGATCGCCAGGCCCTCCAGCCCCACCCCCCGGA 227

RESULT 9

US-09-864-761-20472/c

; Sequence 20472, Application US/09864761

; Patent No. US20020048763A1

GENERAL INFORMATION:

; APPLICANT: Penn, Sharron G.

; APPLICANT: Rank, David R.

; APPLICANT: Hanzel, David K.

; APPLICANT: Chen, Wensheng

; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR

; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY

; FILE REFERENCE: Aecmica-x-1

; CURRENT APPLICATION NUMBER: US/09/864,761

; CURRENT FILING DATE: 2001-05-23

; PRIOR APPLICATION NUMBER: US 60/180,312

; PRIOR FILING DATE: 2000-02-04

; PRIOR APPLICATION NUMBER: US 60/207,456

; PRIOR FILING DATE: 2000-05-26

; PRIOR APPLICATION NUMBER: US 09/632,366

; PRIOR FILING DATE: 2000-08-03

; PRIOR APPLICATION NUMBER: GB 24263.6

; PRIOR FILING DATE: 2000-10-04

; PRIOR APPLICATION NUMBER: US 60/236,359

; PRIOR FILING DATE: 2000-09-27

; PRIOR APPLICATION NUMBER: PCT/US01/00666

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: PCT/US01/00667

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: PCT/US01/00667

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: PCT/US01/00664

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: PCT/US01/00665

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: PCT/US01/00668

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: PCT/US01/00663

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: PCT/US01/00662

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: PCT/US01/00661

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: PCT/US01/00670

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: US 60/234,687

```
/ PRIOR FILING DATE: 2000-09-21
/ PRIOR APPLICATION NUMBER: US 09/608,408
/ PRIOR FILING DATE: 2000-06-30
/ PRIOR APPLICATION NUMBER: US 09/774,203
/ PRIOR FILING DATE: 2001-01-29
/ NUMBER OF SEQ ID NOS: 49117
/ SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
/ SEQ ID NO 20472
/ LENGTH: 473
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ OTHER INFORMATION: MAP TO AP000021.2
/ OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 1.2
/ OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1
/ OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1
/ OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.1
/ OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.97
/ OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.5
/ OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 1.4
/ OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1
/ OTHER INFORMATION: EST HUMAN HIT: R87572.1, EVALUE 0.00e+00
/ OTHER INFORMATION: SWISSPROT HIT: P11308, EVALUE 6.00e-88
/ OTHER INFORMATION: NT HIT: M17254.1, EVALUE 0.00e+00
US-09-864-761-20472

Query Match      11.3%; Score 198.4; DB 9; Length 473;
Best Local Similarity 79.4%; Pred. No. 2.9e-42;
Matches 235; Conservative 0; Mismatches 61; Indels 0; Gaps 0;

QY 547 GCGAGGGGCGAGTCAGTTGGCAGATTTCCTACTGGAGCTGCTGGCAGACCGCGGAAC 606
DB 473 GCGAGTGGCCAGATCCAGCTTTGGCAGTTTGGCAGTTTCTCTGGAGCTCTCTGTCGACAGCTCCAA 414
QY 607 GCCGCTGCATCGCTGGGAGGCGGCGCCAGCGAGTTCAGCTCAAGCTCACCGACCCGACGAG 666
DB 413 TCAGCTGCATCACCTGGAGAGCCAAAGAGAGCAAAACCAACATCACTACGTAAGCTCAGC 354
QY 667 GTGGCGCGACGCTGGGCGGAGCGCAAGAGCAAGCCCAATATGAATCTACGACAGCTAAGT 726
DB 353 GTGGCGCGGCGCTGGGAGAGCGGAAGAGCAAAACCAACATCACTACGTAAGCTCAGC 294
QY 727 CGAGCACTCGCTACTACTACACAAACATCATGAGCAAGGTGACGGCAAGCGCTAC 786
DB 293 CGCGCCCTCGTTACTACTATGACAAGAAACATCATGACCAAGGTCTCCATGGGAAGCGCTAC 234
QY 787 GCCTACCGCTTTGACTTCCAGGCGCTGGCAGAGCTTGCAGCCACCCCGCGCA 842
DB 233 GCCTACAGTTGACTTCCAGGGATCGCCAGGCTTCCAGGCTTCCAGCCCGCGCA 178

RESULT 10
US-10-029-386-25224/c
/ Sequence 25224, Application US/10029386
/ Publication No. US200301947041
/ GENERAL INFORMATION:
/ APPLICANT: Penn, Sharron G.
/ APPLICANT: Rank, David R.
/ APPLICANT: Hanzel, David K.
/ TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
/ FILE REFERENCE: AEOMICA-X-2
/ CURRENT APPLICATION NUMBER: US/10/029,386
/ CURRENT FILING DATE: 2001-12-20
/ NUMBER OF SEQ ID NOS: 34288
/ SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
/ SEQ ID NO 25224
/ LENGTH: 472
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ OTHER INFORMATION: MAP TO AP001731.1
/ OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 0.5
```

```
/ OTHER INFORMATION: NT HIT: g114780450, EVALUE 0.00e+00
/ OTHER INFORMATION: EST HUMAN HIT: R87572.1, EVALUE 0.00e+00
/ OTHER INFORMATION: SWISSPROT HIT: P11308, EVALUE 3.00e-87
US-10-029-386-25224

Query Match      11.3%; Score 197.4; DB 12; Length 472;
Best Local Similarity 79.3%; Pred. No. 5.4e-42;
Matches 234; Conservative 0; Mismatches 61; Indels 0; Gaps 0;

QY 548 GCAGGGGCGAGATCCAGTTGTGGCAGTTTCTACTGGAGCTGCTGGCAGACCGCGGAACG 607
DB 472 GCAGTGGCCAGATCCAGCTTTGGCAGTTTCTCTGGAGCTCTCTGTCGAGCTCCAACT 413
QY 608 CCGGCTGCATCCGCTGGGAGGCGGCGCCAGCGGAGTTCAAGCTCACCGACCCCGACGAGG 667
DB 412 CCAGCTGCATCACCTGGGAGGCGCAACGAGGAGTTCAAGATGACGGATCCCGACGAGG 353
QY 668 TGGCGGAGCGCTGGGCGGAGCGCAAGAGCAAGCCCAATATGAATCTACGACAACTAAGTC 727
DB 352 TGGCGGCGGCTGGGAGAGCGGAAGAGCAAAACCAACATGAATACGATAAGCTCAGCC 293
QY 728 GAGCACTCGCTACTACTACGACAAACATCATGAGCAAGGTGACGGCGCAAGCGCTACG 787
DB 292 GCGCCCTCGTTACTACTATGACAAGAACATCATGACCAAGGTCTCCATGGGAAGCGCTACG 233
QY 788 CCTACCGCTTTGACTTCCAGGCGCTGGCAGAGCTTGCAGCCACCCCGCGCA 842
DB 232 CCTACAAGTTGACTTCCAGGGATCGCCAGGCTTCCAGGCTTCCAGCCCGCGCA 178

RESULT 11
US-10-007-926A-295
/ Sequence 295, Application US/10007926A
/ Publication No. US20030143539A1
/ GENERAL INFORMATION:
/ APPLICANT: BERTUCCI, FRANCOIS
/ APPLICANT: HOULGATTE, REMI
/ APPLICANT: BIRNBAUM, DANIEL
/ APPLICANT: NGUYEN, CATHERINE
/ APPLICANT: VIENS, PATRICE
/ APPLICANT: FERT, VINCENT
/ TITLE OF INVENTION: GENE EXPRESSION PROFILING OF PRIMARY BREAST CARCINOMAS
/ FILE REFERENCE: 1546-R-00
/ CURRENT APPLICATION NUMBER: US/10/007,926A
/ CURRENT FILING DATE: 2001-12-07
/ PRIOR APPLICATION NUMBER: 60/254,090
/ PRIOR FILING DATE: 2000-12-08
/ NUMBER OF SEQ ID NOS: 468
/ SOFTWARE: PatentIn Ver. 2.1
/ SEQ ID NO 295
/ LENGTH: 2957
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ OTHER INFORMATION: friend leukemia virus integration 1 (FLI1)
/ OTHER INFORMATION: gene.
US-10-007-926A-295

Query Match      11.1%; Score 194; DB 12; Length 2957;
Best Local Similarity 79.3%; Pred. No. 1e-40;
Matches 230; Conservative 0; Mismatches 60; Indels 0; Gaps 0;

QY 547 GCGAGGGGCGAGATCCAGTTGTGGCAGTTTCTACTGGAGCTGCTGGCAGACCGCGGAAC 606
DB 1001 GGAAGCGGGCAGATCCAGCTGTGGCAATTCCTCTCGGAGCTGCTCTCCGACAGCGCAAC 1060
QY 607 CCGGCTGCATCCGCTGGGAGGCGGCGCCAGCGGAGTTCAAGCTCACCGACCCCGACGAG 666
DB 1061 GCCAGCTGTATCACTCTGGAGGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1120
QY 667 GTGGCGGCGCTGGGCGGAGCGCAAGAGAGCCCAATATGAATGAGGAGGAGGAGGAGGAGGAG 726
```


Db 1121 GTGGCAGGCGCTGGGGGAGCGGAAAGCAAGCCCAACATGAATTACGACAAGCTGAGC 1180
QY 727 CGAGCACTCGCTACTACTACACAAAAAATCATGCAAGCTGCAAGCGGCAAGCGCTAC 786
Db 1181 CGGGCCCTCCGGTTATTTACTATGATAAAAAATTATGACCAAAAGTGCACGGCAAAAGATAT 1240
QY 787 GCCTACCGCTTTGACTTCCAGGGCCTGGGACACAGGCTTGCACGCCACCAACC 836
Db 1241 GCTTACAAATTTGACTTCCAGGCAATGGCCAGGCTCTGCAGCCACATCC 1290

RESULT 12

US-09-902-772-1
; Sequence 1, Application US/09902772
; Patent No. US20020164739A1
; GENERAL INFORMATION:
; APPLICANT: Chugai Seiyaku Kabushiki Kaisha, 5001, Iwamoto et
; TITLE OF INVENTION: Cell Calcification Suppressing Proteins and Genes of
; FILE REFERENCE: the Proteins
; CURRENT APPLICATION NUMBER: US/09/902,772
; CURRENT FILING DATE: 2001-07-12
; PRIOR APPLICATION NUMBER: US/08/878,177
; PRIOR FILING DATE: 1997-06-18
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 1447
; TYPE: DNA
; ORGANISM: C-11 gene, c-erg gene w/ deletion, chicken DNA
US-09-902-772-1

Query Match 10.3%; Score 180; DB 10; Length 1447;

Best Local Similarity 74.0%; Pred. No. 3.7e-37;
Matches 228; Conservative 0; Mismatches 80; Indels 0; Gaps 0;

QY 529 AGCCCTGGCGGTACAGAAAGCGGAGCGGATCCAGTTGTGGCAGTTTCTACTGGAGCTG 588
Db 879 AGCCGCTCTTGCAAATCCAGGGAGTGGGAGATACAGTATGCGAGTTCTACTGGAGCTT 938
QY 589 GTGGCAGACCGCGCAAGCCGGCTGCATCGGTGGGAGGGCGGCCACGGCAGTTCAAG 648
Db 939 CTGTGGACAGCTCCAACTCCAACTGCACTCGCTGGGAGGGGACAAATGGGGAGTTCAAG 998
QY 649 CTCACCGACCCGACGAGGTGGCGACCGCTGGGCGGAGCGCAAGCAAGCCCAATATG 708
Db 999 ATGACAGACCTTGATGAGTGGCTCGGCTGGGAGAGAGGAAAGCAAACTTAACATG 1058
QY 709 AACTAGCAACGCTAAGTGCAGCACTGCGCTACTACTAGCAAAAAACATCATGAGCAAG 768
Db 1059 AACTATGACAAACTCAGCGGTGCACTTCGCTACTACTATGACAAAAATATTATGACTAAA 1118
QY 769 GTGCAGCGCAAGCGCTACCGCTTACCGCTTTGACTTCCAGGGCTGGCAGAGCTTGGCAG 828
Db 1119 GTTCATGTAACGCTATGCTTACAAATTTGATTTTCCACGGAATCGCTCAGGCCCTCCAG 1178
QY 829 CCACCAACC 836
Db 1179 CCTCACCC 1186

RESULT 13

US-09-902-772-3
; Sequence 3, Application US/09902772
; Patent No. US20020164739A1
; GENERAL INFORMATION:
; APPLICANT: Chugai Seiyaku Kabushiki Kaisha, 5001, Iwamoto et
; TITLE OF INVENTION: Cell Calcification Suppressing Proteins and Genes of
; FILE REFERENCE: the Proteins
; CURRENT APPLICATION NUMBER: US/09/902,772
; CURRENT FILING DATE: 2001-07-12
; PRIOR APPLICATION NUMBER: US/08/878,177

; PRIOR FILING DATE: 1997-06-18
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 1528
; TYPE: DNA
; ORGANISM: c-erg gene, chicken DNA
US-09-902-772-3

Query Match 10.3%; Score 180; DB 10; Length 1528;

Best Local Similarity 74.0%; Pred. No. 3.8e-37;
Matches 228; Conservative 0; Mismatches 80; Indels 0; Gaps 0;

QY 529 AGCCCTGGCGGTACAGAAAGCGGAGCGGATCCAGTTGTGGCAGTTTCTACTGGAGCTG 588
Db 960 AGCCGCTCTTGCAAATCCAGGGAGTGGGAGATACAGTATGCGAGTTCTACTGGAGCTT 1019
QY 589 GTGGCAGACCGCGCAAGCCGGCTGCATCGGTGGGAGGGCGGCCACGGCAGTTTCAAG 648
Db 1020 CTGTGGACAGCTCCAACTCCAACTGCACTCGCTGGGAGGGGACAAATGGGGAGTTCAAG 1079
QY 649 CTCACCGACCCGACGAGGTGGCGGAGCGCTGGGCGGAGCGCAAGCAAGCCCAATATG 708
Db 1080 ATGACAGACCTTGATGAAGTGGCTCGGCTGGGAGAGAGGAAAGCAAACTTAACATG 1139
QY 709 AACTAGCAACGCTAAGTGCAGCACTGCGCTACTACTAGCAAAAAACATCATGAGCAAG 768
Db 1140 AACTATGACAAACTCAGCGGTGCACTTCGCTACTACTATGACAAAAATATTATGACTAAA 1199
QY 769 GTGCAGCGCAAGCGCTACCGCTTACCGCTTTCGACTTCCAGGGCTGGCAGAGCTTGGCCAG 828
Db 1200 GTTCATGTAACGCTATGCTTACAAATTTGATTTCCACGGAATCGCTCAGGCCCTCCAG 1259
QY 829 CCACCAACC 836
Db 1260 CCTCACCC 1267

RESULT 14

US-10-210-120-41
; Sequence 41, Application US/10210120
; Publication No. US2003017536A1
; GENERAL INFORMATION:
; APPLICANT: Chinnaiyan, Arul M.
; APPLICANT: Rubin, Mark A.
; APPLICANT: Sreekumar, Arun
; TITLE OF INVENTION: Expression Profile of Prostate Cancer
; FILE REFERENCE: UM-07221
; CURRENT APPLICATION NUMBER: US/10/210,120
; CURRENT FILING DATE: 2002-08-01
; PRIOR APPLICATION NUMBER: US 60/309,581
; PRIOR FILING DATE: 2001-08-02
; PRIOR APPLICATION NUMBER: US 60/334,468
; PRIOR FILING DATE: 2001-11-15
; NUMBER OF SEQ ID NOS: 123
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 41
; LENGTH: 441
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-210-120-41

Query Match 7.9%; Score 138.8; DB 12; Length 441;

Best Local Similarity 69.6%; Pred. No. 1.8e-26;
Matches 188; Conservative 0; Mismatches 82; Indels 0; Gaps 0;

QY 544 AAAGCAGCGGCGGAGATCCAGTTGTGGCAGTTTCTACTGGAGCTCTGGCAGACCGCGG 603
Db 103 ACAGAAAGTGACCTATTTCAGCTGGCAGTTTCTCTCTGGAGCTCTATCAGCAAAATCC 162
QY 604 AACGCGGCTGCATCGCTGGAGGGCGGCCACGGCGAGTTCAAGCTCACCGACCCCGAC 663
Db 163 TGCCAGTCATTTCATCAGCTGGAGCTGGAGCGGAGTTTAAGCTCGC-CACGCCCGAT 222

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 26, 2003, 13:20:22 ; Search time 114 Seconds
(without alignments)
6783.362 Million cell updates/sec

Title: US-10-027-859-1
Perfect score: 1752
Sequence: 1 ttgttttaacaacatgtt.....aaaaaaaaaaaaaaaaaaaaa 1752

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA.*
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq.*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq.*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq.*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq.*
5: /cgn2_6/ptodata/1/ina/PTUS_COMB.seq.*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Length	DB ID	Description
1	1752	100.0	1752 3	US-09-360-779-1
2	1752	100.0	1752 4	US-09-435-335-1
3	233.4	13.3	1754 4	US-09-604-978-1
4	233.4	13.3	1754 4	US-09-604-728-1
5	223	12.7	2532 4	US-09-620-312D-869
6	194	11.1	2938 2	US-08-343-443B-3
7	180	10.3	1447 3	US-08-878-177-1
8	180	10.3	1528 3	US-08-878-177-3
9	138.8	7.9	2268 3	US-09-344-579-1
10	134.8	7.7	2667 2	US-08-469-412A-1
11	134.8	7.7	2667 3	US-09-021-715-1
12	121	6.9	1734 3	US-09-146-969-5
13	121	6.9	1782 4	US-09-604-978-2
14	121	6.9	1782 4	US-09-604-728-2
15	115.4	6.6	1604 1	US-08-306-691B-43
16	115.4	6.6	1604 5	PCT-US93-06251-9
17	114.8	6.6	2064 3	US-08-875-944B-1
18	114.8	6.6	2064 3	US-09-116-049-3
19	114.8	6.6	2064 4	US-09-602-868A-1
20	114.8	6.6	2064 4	US-09-884-363-3
21	110	6.3	2265 2	US-09-213-767-1
22	108.8	6.2	2410 2	US-08-780-835B-1
23	108.8	6.2	2410 3	US-09-303-268-1
24	108.8	6.2	2410 3	US-09-116-049-1
25	108.8	6.2	2410 4	US-09-884-363-1
26	103.4	5.9	2544 2	US-08-469-412A-6
27	103.4	5.9	2544 3	US-09-021-715-6

28 97 5.5 7218 1 US-08-232-463-14 Sequence 14, Appli
29 90.4 5.2 1933 4 US-09-920-759-3 Sequence 3, Appli
30 90.4 5.2 1976 4 US-09-920-759-10 Sequence 10, Appli
31 86 4.9 665 4 US-09-920-759-11 Sequence 11, Appli
32 74 4.2 15788 4 US-09-920-759-13 Sequence 13, Appli
33 68 3.9 1894 4 US-09-570-593-1 Sequence 1, Appli
34 68 3.9 1905 3 US-09-055-113-2 Sequence 2, Appli
35 68 3.9 3317 4 US-09-570-593-12 Sequence 12, Appli
36 66.6 3.8 328 2 US-08-343-443B-5 Sequence 5, Appli
37 66.6 3.8 1907 4 US-09-300-958A-27 Sequence 27, Appli
38 66.6 3.8 1907 4 US-09-570-593-4 Sequence 4, Appli
39 65.8 3.8 65042 4 US-09-784-316-3 Sequence 3, Appli
40 65 3.7 1920 1 US-08-746-789A-1 Sequence 1, Appli
41 62.6 3.6 1364 1 US-08-306-691B-50 Sequence 50, Appli
42 62.6 3.6 1364 5 PCT-US93-06251-65 Sequence 65, Appli
43 60.6 3.5 320 3 US-09-165-264-14 Sequence 14, Appli
44 60.4 3.4 502 4 US-09-389-681-282 Sequence 282, App
45 60.4 3.4 502 4 US-09-620-405B-282 Sequence 282, App

ALIGNMENTS

RESULT 1
US-09-360-779-1
; Sequence 1, Application US/09360779
; Patent No. 6268216
; GENERAL INFORMATION:
; APPLICANT: Denaris, Evan S.
; APPLICANT: Hendricks, Timothy J.
; TITLE OF INVENTION: Reagents and Methods for the Screening of Compounds
; TITLE OF INVENTION: Useful in the Treatment of Neurological Diseases
; FILE REFERENCE: CASE-03828
; CURRENT APPLICATION NUMBER: US/09/360,779
; CURRENT FILING DATE: 1998-07-26
; EARLIER APPLICATION NUMBER: 60/094,264
; EARLIER FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 1752
; TYPE: DNA
; ORGANISM: Rattus norvegicus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (112)..(1131)
US-09-360-779-1

Query Match 100.0%; Score 1752; DB 3; Length 1752;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1752; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 TTGTTTTTAAACAATGTTTATTAGAAAAGTAAATAATTTGCATAGTCTTAGTACTTG 60
Db 1 TTGTTTTTAAACAATGTTTATTAGAAAAGTAAATAATTTGCATAGTCTTAGTACTTG 60
Qy 61 AACATCAAGTGTATTTCATGAACCGTGAATCTTTCATGTAAACAGTTCATAGTGAAGAC 120
Db 61 AACATCAAGTGTATTTCATGAACCGTGAATCTTTCATGTAAACAGTTCATAGTGAAGAC 120
Qy 121 CCAGTGGCGCTCTCTCTGGGGGAGAGGTTTCCAGCCCCCCCCCCTCAGCCCCCATCCC 180
Db 121 CCAGTGGCGCTCTCTCTGGGGGAGAGGTTTCCAGCCCCCCCCCCTCAGCCCCCATCCC 180
Qy 181 CTACAGCTACTCTCTCCAGTACACCGGCACCGGGATGGGTGGGATGAGCTCCAGGAC 240
Db 181 CTACAGCTACTCTCTCCAGTACACCGGCACCGGGATGGGTGGGATGAGCTCCAGGAC 240
Qy 241 CCCCTCTCTCTCACCACACCCCTGGCTCCCGCTCCCGCCAGGCTTGTCCGGACCCGGCG 300
Db 241 CCCCTCTCTCTCACCACACCCCTGGCTCCCGCTCCCGCCAGGCTTGTCCGGACCCGGCG 300

TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-604-978-1

Query Match 13.3%; Score 233.4; DB 4; Length 1754;
Best Local Similarity 99.6%; Pred. No. 1.3e-45;
Matches 234; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 TTGTTTAAACAACATGTTTATTAGAAAAGTAAAAATATTGTCATAGGCTTTAGTACTTG 60
Db 1733 TTGTTTAAACAACATGTTTATTAGAAAAGTAAAAATATTGTCATAGGCTTTAGTACTTG 1674

Qy 61 AACATCAAGTGTATTTCATGAACCGTGAAGGTTCCAGCCCCCCTCAGCCCCCATCCC 120
Db 1673 AACATCAAGTGTATTTCATGAACCGTGAAGGTTCCAGCCCCCCTCAGCCCCCATCCC 1614

Qy 121 CCAGGTGGCGCTCTCTCTGGGGGAGAGGTTCCAGCCCCCCTCAGCCCCCATCCC 180
Db 1613 CCAGGTGGCGCTCTCTCTGGGGGAGAGGTTCCAGCCCCCCTCAGCCCCCATCCC 1554

Qy 181 CTCACAGCTCACTCTCCATACACCGGATGAGGTTCCAGCCCCCCTCAGCCCCCATCCC 235
Db 1553 CTCACAGTTCACCTCTCCATACACCGGATGAGGTTCCAGCCCCCCTCAGCCCCCATCCC 1499

RESULT 4

US-09-604-728-1/c
Sequence 1, Application US/09604728
Patent No. 6555667
GENERAL INFORMATION:
APPLICANT: Elnat, Paz
TITLE OF INVENTION: HPOXIA-REGULATED GENES
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESSEE: KOHN & ASSOCIATES
STREET: 30500 No. 6555667thwestern Hwy., Suite 401
CITY: Farmington Hills
STATE: Michigan
COUNTRY: U.S.
ZIP: 48334
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/604,728
FILING DATE: 28-Jun-2000
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/09/138,112
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Kohn, Kenneth I.
REGISTRATION NUMBER: 30,955
REFERENCE/DOCKET NUMBER: 0168.00034
TELECOMMUNICATION INFORMATION:
TELEPHONE: (248) 539-5050
TELEFAX: (248) 5395055
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1754 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
SEQUENCE DESCRIPTION: SEQ ID NO: 1:

US-09-604-728-1

Query Match 13.3%; Score 233.4; DB 4; Length 1754;
Best Local Similarity 99.6%; Pred. No. 1.3e-45;
Matches 234; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 TTGTTTAAACAACATGTTTATTAGAAAAGTAAAAATATTGTCATAGGCTTTAGTACTTG 60
Db 1733 TTGTTTAAACAACATGTTTATTAGAAAAGTAAAAATATTGTCATAGGCTTTAGTACTTG 1674

Qy 61 AACATCAAGTGTATTTCATGAACCGTGAAGGTTCCAGCCCCCCTCAGCCCCCATCCC 120
Db 1673 AACATCAAGTGTATTTCATGAACCGTGAAGGTTCCAGCCCCCCTCAGCCCCCATCCC 1614

Qy 121 CCAGGTGGCGCTCTCTCTGGGGGAGAGGTTCCAGCCCCCCTCAGCCCCCATCCC 180
Db 1613 CCAGGTGGCGCTCTCTCTGGGGGAGAGGTTCCAGCCCCCCTCAGCCCCCATCCC 1554

Qy 181 CTCACAGCTCACTCTCCATACACCGGATGAGGTTCCAGCCCCCCTCAGCCCCCATCCC 235
Db 1553 CTCACAGTTCACCTCTCCATACACCGGATGAGGTTCCAGCCCCCCTCAGCCCCCATCCC 1499

RESULT 5

US-09-620-312D-869
Sequence 869, Application US/09620312D
Patent No. 6569662
GENERAL INFORMATION:
APPLICANT: Tang, Y. Tom
APPLICANT: Liu, Chenghua
APPLICANT: Asundi, Vinod
APPLICANT: Zhang, Jie
APPLICANT: Ren, Felyan
APPLICANT: Chen, Rui-hong
APPLICANT: Zhao, Qing A.
APPLICANT: Wehrman, Tom
APPLICANT: Xue, Aidong J.
APPLICANT: Yang, Yonghong
APPLICANT: Wang, Jian-Rui
APPLICANT: Zhou, Ping
APPLICANT: Ma, Yunding
APPLICANT: Wang, Dunrui
APPLICANT: Wang, Zhiwei
APPLICANT: John Tillinghast
APPLICANT: Drmanac, Radoje T.
TITLE OF INVENTION: No. 6569662el Nucleic Acids and
FILE REFERENCE: Polypeptides
FILE REFERENCE: 784CIP2B
CURRENT APPLICATION NUMBER: US/09/620,312D
CURRENT FILING DATE: 2000-07-19
PRIOR APPLICATION NUMBER: 09/552,317
PRIOR FILING DATE: 2000-04-25
PRIOR APPLICATION NUMBER: 09/488,725
PRIOR FILING DATE: 2000-01-21
NUMBER OF SEQ ID NOS: 1105
SOFTWARE: dt_FL_genes Version 1.0
SEQ ID NO 869
LENGTH: 2532
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (1)..(2532)
US-09-620-312D-869

Query Match 12.7%; Score 223; DB 4; Length 2532;
Best Local Similarity 72.2%; Pred. No. 4.3e-43;
Matches 366; Conservative 0; Mismatches 70; Indels 71; Gaps 3;

Qy 546 AGGCAGCGGCGAGATCCAGATTGTGGCAGGTTCTTCTATGAGAGTGTGGCAGACCGCGCGAA 605
Db 2096 ATGCAGCGGCGAGATCCAGCTGTGGCAGTTTCTGTGTGAGCTGTCTGGCTGACCGCGCGAA 2155


```

QY 606 CCGCGGCTGCTATCGGTGGAGGGGCGGCCACCGCGAGTTCAAGCTACCGACCCCGACGA 665
Db |||||
QY 2156 CG-CGGCTCATCGCTGGAGGGGGGTCAAGCTCAAGCTCAGGACCCGACGA 2214
Db |||||
QY 666 GGTGGCGGACCTGGGCGGCGGAGAGAGAGCCCAATATGAAGTCAAGTCAAGCTAAG 725
Db |||||
QY 2215 GGTGGCGGCGGTGGGCGGCGGAGAGAGAGCCCAATATGAAGTCAAGTCAAGCTAAG 2274
Db |||||
QY 726 TCGAGCACTGCGCTACTACTACGACAAACATCATGAGCAAGTGCACGGCAAGCGCTA 785
Db |||||
QY 2275 CCGCGCCCTGCTAT-CTACTACGACAGACATCATGAGCAAGTGCATGTTAAGCGCTA 2333
Db |||||
QY 786 CCGCTACCGCTTGTACTTCCAGGCGCTGCGACAGGTTGCCAGCCACCCGCGACGC 845
Db |||||
QY 2334 CCGCTACCGCTTGTACTTCCAGGCGCTGCGACAGGTTGCCAGCCACCCGCGACGC 2361
Db |||||
QY 846 CCACCGCGCGCTGCGCGCGCGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCTTACAA 905
Db |||||
QY 2362 -----GGCAGGAGCGCGCGCGCTTACAA 2384
Db |||||
QY 906 GCTCCCGGCTGCTGTGGCTTCCACTGCTTCCCGCGCTCTCCAACTCAAGCTTATGCG 965
Db |||||
QY 2385 GCTCCCGCGCGCTTCCAGGCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCTTCCAACTCAAGCTTATGCG 2444
Db |||||
QY 966 AGCCTGCGCGCGCTGCGCGCGCGCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCTTCCAACTCAAGCTTATGCG 1025
Db |||||
QY 2445 CCGCTGCGTGGAGTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTTCCAACTCAAGCTTATGCG 1052
Db |||||
QY 1026 CCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTTCCAACTCAAGCTTATGCG 1052
Db |||||
QY 2505 CCGCTGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTTCCAACTCAAGCTTATGCG 2531
Db |||||

```

```

RESULT 6
US-08-343-443B-3
; Sequence 3, Application US/08343443B
; Patent No. 5968734
; GENERAL INFORMATION:
; APPLICANT: Aurias, Alain
; APPLICANT: Delattre, Olivier
; APPLICANT: Desmaze, Chantal
; APPLICANT: Melot, Thomas
; APPLICANT: Peter, Martine
; APPLICANT: Ploougaestel, Beatrice
; APPLICANT: Thomas, Gilles
; APPLICANT: Zucman, Jessica
; TITLE OF INVENTION: NUCLEIC ACID CORRESPONDING TO A GENE OF
; TITLE OF INVENTION: CHROMOSOME 22 INVOLVED IN RECURRENT CHROMOSOMAL
; TITLE OF INVENTION: TRANSLATIONS ASSOCIATED WITH THE DEVELOPMENT OF CANCEROUS
; TITLE OF INVENTION: TUMORS, AND NUCLEIC ACIDS OF FUSION RESULTING FROM SAID
; NUMBER OF SEQUENCES: 129
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Weiser & Associates
; STREET: 230 South Fifteenth Street
; CITY: Philadelphia
; STATE: PA
; COUNTRY: USA
; ZIP: 19102
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: AEDIT 1.0 DOS text editor
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/343,443B
; FILING DATE: 18-NOV-1994
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/FR93/00494
; FILING DATE: 19-MAY-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: FR 92/06123

```

```

; FILING DATE: 20-MAY-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Weiser, Gerard J.
; REGISTRATION NUMBER: 19,763
; REFERENCE/DOCKET NUMBER: 989.6121P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 215-875-8383
; TELEFAX: 215-875-8394
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2938 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 143..1498
; US 08-343-443B-3

Query Match 11.1%; Score 194; DB 2; Length 2938;
Best Local Similarity 79.3%; Pred. No. 3.1e-36;
Matches 230; Conservative 0; Mismatches 60; Indels 0; Gaps 0;

QY 547 GGCAGCGGCGAGATCCAGTTGTGGCAGTTTCTACTGAGAGTGTCTGGCAGACCGCGCAAC 606
Db |||||
QY 971 GGAAGCGGCGAGATCCAGTTGTGGCAATTCCTCTGGAGTGTCTCTCCGACAGCGCAAC 1030
Db |||||
QY 607 GCGGCTCATCGTGGTGGAGGGCGGCCAGCGGAGTTCAAGTCCACGACCCCGACGAG 666
Db |||||
QY 1031 GCCAGCTGTATCACCTGGAGGGGACCAACCGGGAGTTCAAAATGACGACCCCGATGAG 1090
Db |||||
QY 667 GTGGCGGACGCTGGGCGGAGCGCAAGAGCAAGCCCAATATGAATACGACAAAGCTTAAGT 726
Db |||||
QY 1091 GTGGCCAGGCGCTGGGCGGAGCGGGAAGAGCCCAACATGAATACGACAAAGCTGAGC 1150
Db |||||
QY 727 CGAGCACTGCGCTACTACTACGACAAAACATCATGAGCAAGTGCACGGCAAGCGCTAC 786
Db |||||
QY 1151 CGGGCCCTCCGTTATTACTATGATAAAAACATTTATGACAAAAGTGCAGGCAAAAGATAT 1210
Db |||||
QY 787 GCCTACCGCTTTGACTTCCAGGGCTGCGACAGGTTGCCAGCCACCACC 836
Db |||||
QY 1211 GCTTACAAAATTGACTTCCAGCGCATTCGCCAGGCTCTGCGACCCACATCC 1260
Db |||||

RESULT 7
US-08-878-177-1
; Sequence 1, Application US/08878177
; Patent No. 6294354
; GENERAL INFORMATION:
; APPLICANT: Chugai Seiyaku Kabushiki Kaisha, 5001, Iwamoto et
; TITLE OF INVENTION: Cell Calcification Suppressing Proteins and Genes of
; TITLE OF INVENTION: the Proteins
; FILE REFERENCE: chugai seiyaku kabushiki kaisha 5001
; CURRENT APPLICATION NUMBER: US/08/878,177
; CURRENT FILING DATE: 1997-06-18
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 1447
; TYPE: DNA
; ORGANISM: C-11 gene, c-erg gene w/ deletion, chicken DNA
; US-08-878-177-1

Query Match 10.3%; Score 180; DB 3; Length 1447;
Best Local Similarity 74.0%; Pred. No. 4.5e-33;
Matches 228; Conservative 0; Mismatches 80; Indels 0; Gaps 0;

QY 529 AGCCCTCGGTACAGAAAGGCGGAGATCCAGTTGTGGCAGTTTCTACTGAGCTG 588
Db |||||
QY 879 AGCCGCTTTCGAATCCAGGAGTGGGAGATACAGCTATGCGAGTTCTCTACTGAGCTT 938
Db |||||
QY 589 CTGGCAGACCGCGCAACGCCGCTGTCATCGCTGGGAGGGCGGCGGAGTTCAAG 648
Db |||||

```

Db 939 CTGTCGACAGCTCCAACTCCAATGCATCACCTGGGAGGCGCAAAATGGGAGTTCAAG 998
 QY 649 CTCACGACCCGACGAGGTGGCGGACGCTGGGGCGAGCGCAAGAGCAAGCCCAATATG 708
 Db 999 ATCAGAGACCTGATGAAGTGGCTGGCGGTGGGAGAGAGGAAAGCAACCTAATCATG 1058
 QY 709 AACTAGCACAAGCTAAGTCGAGCACTGCGCTACTTACTACGACAAACATCATGAGCAAG 768
 Db 1059 AACTATGACAACTCAGCGTGCACTTCTGCTACTATGACAAAATATATGACTAAA 1118
 QY 769 GTCCAGCGCAAGCGCTACGCTACCGCTTACCGCTTTGACTTCCAGGCGCTGGCACAGGCTTGCCAG 828
 Db 1119 GTTCATGGTAAACGCTATGCTCAAAATTTGATTTCCACGGAATCGCTCAGGCCCTCCAG 1178
 QY 829 CCACCACC 836
 Db 1179 CCTCACCC 1186

RESULT 8
 US-08-878-177-3
 ; Sequence 3, Application US/08878177
 ; Patent No. 6294354
 ; GENERAL INFORMATION:
 ; APPLICANT: Chugai Seiyaku Kabushiki Kaisha, 5001, Iwamoto et
 ; TITLE OF INVENTION: Cell Calcification Suppressing Proteins and Genes of
 ; TITLE OF INVENTION: the Proteins
 ; FILE REFERENCE: Chugai Seiyaku Kabushiki Kaisha 5001
 ; CURRENT APPLICATION NUMBER: US/08/878,177
 ; CURRENT FILING DATE: 1997-06-18
 ; NUMBER OF SEQ ID NOS: 7
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 3
 ; LENGTH: 1528
 ; TYPE: DNA
 ; ORGANISM: c-erg gene, chicken DNA
 ; US-08-878-177-3

Query Match 10.3%; Score 180; DB 3; Length 1528;
 Best Local Similarity 74.0%; Pred. No. 4.6e-33;
 Matches 228; Conservative 0; Mismatches 80; Indels 0; Gaps 0;
 QY 529 AGCCCTCGGTACAGAAAGCAGCGGCGAGATCCAGTTGTGGCAGTTTCTACTGGAGCTG 588
 Db 960 AGCCGCTTTGCAATCCAGGAGTGGCGAGATACAGCTATGCGAGTTCTACTGGAGCTT 1019
 QY 589 CTGGCAGACCGCGCAACCGCGCTGCATCGCTGGGAGGCGGCGCGAGTTTCAAG 648
 Db 1020 CTGTCGGACAGCTCCAACTCCAATCGCATCACCTGGGAGGCGCAAAATGGGAGTTCAAG 1079
 QY 649 CTCACCGACCCGACGAGGTGGCGGCGAGCTGGGCGGCGCAAGAGCAAGCCCAATATG 708
 Db 1080 ATGACAGACCTTGATGAGTGGCTCGCGGTGGGAGAGAGGAAAGCAAACTAATATG 1139
 QY 709 AACTAGCACAAGCTAAGTCGAGCACTGCGCTACTTACTACGACAAACATCATGAGCAAG 768
 Db 1140 AACTATGACAACTCAGCGTGCACTTCTGCTACTATGACAAAATATATGACTAAA 1199
 QY 769 GTGCAGCGCAAGCGCTACCGCTTACCGCTTTGACTTCCAGGCGCTGGCACAGGCTTGCCAG 828
 Db 1200 GTTCATGGTAAACGCTATGCTCAAAATTTGATTTCCACGGAATCGCTCAGGCCCTCCAG 1259
 QY 829 CCACCACC 836
 Db 1260 CCTCACCC 1267

RESULT 9
 US-09-344-579-1
 ; Sequence 1, Application US/09344579
 ; Patent No. 6054316
 ; GENERAL INFORMATION:
 ; APPLICANT: Brenda F. Baker

; APPLICANT: Lex M. Cowser
 ; TITLE OF INVENTION: ANTISENSE MODULATION OF ETS-2 EXPRESSION
 ; FILE REFERENCE: RTS-0063
 ; CURRENT APPLICATION NUMBER: US/09/344,579
 ; CURRENT FILING DATE: 1999-06-25
 ; NUMBER OF SEQ ID NOS: 47
 ; SEQ ID NO 1
 ; LENGTH: 2268
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (291)..(1700)
 ; US-09-344-579-1

Query Match 7.9%; Score 138.8; DB 3; Length 2268;
 Best Local Similarity 69.6%; Pred. No. 2.8e-23;
 Matches 188; Conservative 0; Mismatches 82; Indels 0; Gaps 0;
 QY 544 AAAGCGAGCGGCGAGATCCAGTTGGCAGTTTCTACTGGAGCTGCTGGCAGACCGCGG 603
 Db 1362 ACAGGAAGTGGACCTATTTCAGCTGTGGCAGTTTCTCTGGAGCTGCTATCAGACAAATCC 1421
 QY 604 AACGCGCGTGCATCGCTGGGAGGCGCGCCAGCGGAGTTCAAGCTCACCAGCCCGAC 663
 Db 1422 TGCCAGTCAATCATCAGCTGGACTGGAGACGGATGGAGCTCGCCGACCCCGAT 1481
 QY 664 GAGGTGGCGCGAGCTGGGGCGAGCGCAAGAGCAAGCCCAATATGAATCAACAAAGCTA 723
 Db 1482 GAGGTGGCGCGCGGTGGGGAAGAGGAAATAAGCCCAAGATGAATACGAGAAGCTG 1541
 QY 724 AGTCAGCACTCGCTACTACTACGACAAACATCATGAGCAAGTGCACGCGCAAGCGC 783
 Db 1542 AGCCGCGGCTTACGCTACTATTACGACAAACATCATCCAAAGACGTCGGGGAAGCGC 1601
 QY 784 TACGCTACCGCTTTGACTTCCAGGCGCTG 813
 Db 1602 TACGTGTACCGCTTCTGTGCGACCTCCAG 1631

RESULT 10
 US-08-469-412A-1
 ; Sequence 1, Application US/08469412A
 ; Patent No. 5856125
 ; GENERAL INFORMATION:
 ; APPLICANT: Mavrothalassitis, George J.
 ; APPLICANT: Blair, Donald G.
 ; APPLICANT: Fisher, Robert J.
 ; APPLICANT: Beal Jr., Gregory J.
 ; APPLICANT: Athanasios, Metropi A.
 ; APPLICANT: Sgouras, Dionysios N.
 ; TITLE OF INVENTION: The ERF Genetic Locus and Its Products
 ; NUMBER OF SEQUENCES: 16
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSER: Townsend and Townsend and Crew LLP
 ; STREET: Two Embarcadero Center, Eighth Floor
 ; CITY: San Francisco
 ; STATE: California
 ; COUNTRY: USA
 ; ZIP: 94111-3834
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/469,412A
 ; FILING DATE: 05-JUN-1995
 ; CLASSIFICATION: 435
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Garrett-Wackowski, Eugenia
 ; REGISTRATION NUMBER: 37,330
 ; REFERENCE/DOCKET NUMBER: 015280-229000

```
;
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2667 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 123..1769
; OTHER INFORMATION: /note= "human ERF (ETS2 Repressor
US-08-469-412A-1

Query Match          7.7%; Score 134.8; DB 2; Length 2667;
Best Local Similarity 68.2%; Pred. No. 2.7e-22;
Matches 187; Conservative 0; Mismatches 87; Indels 0; Gaps 0;

Qy 547 GGCAGCGGGCAGATCCAGTTGTGGCAGTTTCTACTGGAGCTGCTGGCAGACCGCGCGAAC 606
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 189 GGCTCAAGCGCAGATCCAGCTGTGGCAGTTTATCTCTGGAGCTGCTGCGGAAGGAGGTAC 248

Qy 607 GCCGCTGCATCGCTGGAGGGCGGCCACGGCGAGTTCAAGCTCACCGACCCCGACGAG 666
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 249 CAGGGCGCTATTGCTTGGCAGGGGAGCTACGGGGAATTTCGTCATCAAAAGACCTGATGAG 308

Qy 667 GTGGCGCCAGCTGGGGGAGCGCAAGAGCAGCCCAATATGACTAGCAAGCTAAGT 726
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 309 GTGGCCCGCTGTGGGGGCTTCGCAAGTGCAGGCCGATTCGCAAGACCAAGGGGAAACGGTTC 368

Qy 727 CGAGCAGCTGCGCTACTACTACGACAAAAACATCATGAGCAAGGTGCACGGCAAGCGTAC 786
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 369 CGGGCCCTGCGCTATTACTATAACAGCGCATTCGCAAGACCAAGGGGAAACGGTTC 428

Qy 787 GCCTACCGCTTTGACTTCCAGGGCTGGCAGG 820
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 429 ACCTACAAGTTCAATTTCAACAAACTGGTGCTGG 462

RESULT 11
US-09-021-715-1
; Sequence 1, Application US/09021715
; Patent No. 6194547
; GENERAL INFORMATION:
; APPLICANT: Mavrothalassitis, George J.
; Blair, Donald G.
; Fisher, Robert J.
; Beal Jr., Gregory J.
; Athanasidou, Meropi A.
; Sgouras, Dionysios N.
; TITLE OF INVENTION: The ERF Genetic Locus and Its Products
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/021,715
; FILING DATE: 10-Feb-1998
; CLASSIFICATION: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Garrett-Wackowski, Eugenia
```

```
;
; REGISTRATION NUMBER: 37,330
; REFERENCE/DOCKET NUMBER: 015280-229000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2667 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 123..1769
; OTHER INFORMATION: /note= "human ERF (ETS2 Repressor
US-09-021-715-1

Query Match          7.7%; Score 134.8; DB 3; Length 2667;
Best Local Similarity 68.2%; Pred. No. 2.7e-22;
Matches 187; Conservative 0; Mismatches 87; Indels 0; Gaps 0;

Qy 547 GGCAGCGGGCAGATCCAGTTGTGGCAGTTTCTACTGGAGCTGCTGGCAGACCGCGCGAAC 606
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 189 GGCTCAAGCGCAGATCCAGCTGTGGCAGTTTATCTCTGGAGCTGCTGCGGAAGGAGGTAC 248

Qy 607 GCCGCTGCATCGCTGGAGGGCGGCCACGGCGAGTTCAAGCTCACCGACCCCGACGAG 666
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 249 CAGGGCGCTATTGCTTGGCAGGGGAGCTACGGGGAATTTCGTCATCAAAAGACCTGATGAG 308

Qy 667 GTGGCGCCAGCTGGGGGAGCGCAAGAGCAGCCCAATATGACTAGCAAGCTAAGT 726
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 309 GTGGCCCGCTGTGGGGGCTTCGCAAGTGCAGGCCGATTCGCAAGACCAAGGGGAAACGGTTC 368

Qy 727 CGAGCAGCTGCGCTACTACTACGACAAAAACATCATGAGCAAGGTGCACGGCAAGCGTAC 786
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 369 CGGGCCCTGCGCTATTACTATAACAGCGCATTCGCAAGACCAAGGGGAAACGGTTC 428

Qy 787 GCCTACCGCTTTGACTTCCAGGGCTGGCAGG 820
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 429 ACCTACAAGTTCAATTTCAACAAACTGGTGCTGG 462

RESULT 12
US-09-146-969-5/C
; Sequence 5, Application US/09146969
; Patent No. 6228585
; GENERAL INFORMATION:
; APPLICANT: Dieckgraefe, Brian K.
; TITLE OF INVENTION: Gene Markers for Chronic Mucosal Injury
; FILE REFERENCE: 04255.75314
; CURRENT APPLICATION NUMBER: US/09/146,969
; CURRENT FILING DATE: 1998-09-04
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
; LENGTH: 1734
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-146-969-5

Query Match          6.9%; Score 121; DB 3; Length 1734;
Best Local Similarity 83.4%; Pred. No. 3.9e-19;
Matches 176; Conservative 0; Mismatches 25; Indels 10; Gaps 3;

Qy 1 TTTGTTTTTAAACAACATGTTTATTAGAAAAGT-AAAAATATTGCATAGTCTTAGTACTT 59
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1725 TTGTTTTTAAACAACATGTTTATTAGAAAAGT-AAAAATATTGCATAGTCTTAGTACTT 1666

Qy 60 GAACATCAAGTGTTATTCATGAACCGTGAG-----TATCTTCATGTAACAGTCTTAGATG 114
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

Db 1665 GAACATCAAGTGATTTATGAAACAGTGAGTATCTTATCTTATGTAAGTAAACAGTTCTAGATG 1606
QY 115 GAAGACCCAGGTGGCTCTCTGCGGGAGAGGTTCCAGCCCCCAGCCCCCTAGCCC 174
Db 1605 GAAGACCCAGATGGCACTCTCTCCCGGGAGGGGTTCCAGCCCCCAGCCCCCTCTC-----AGCC 1550
QY 175 CATCCCTCACAGCTCACTCTCTCCAGTACAC 205
Db 1549 CTTCCCTGCCAGCTCAACTCTGCAGTACAC 1519

RESULT 13

US-09-604-978-2/c
; Sequence 2, Application US/09604978
; Patent No. 6455674
; GENERAL INFORMATION:
; APPLICANT: Einat, Paz
; Skaliter, Rami
; TITLE OF INVENTION: HYPOXIA-REGULATED GENES
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: KOHN & ASSOCIATES
; STREET: 30500 No. 6455674thwestern Hwy., Suite 401
; CITY: Farmington Hills
; STATE: Michigan
; COUNTRY: U.S.
; ZIP: 48334

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/604,978
FILING DATE: 28-Jun-2000
CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 09/138,112
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Kohn, Kenneth I.
REGISTRATION NUMBER: 30,955
REFERENCE/DOCKET NUMBER: 0168.00034
TELECOMMUNICATION INFORMATION:
TELEPHONE: (248) 539-5050
TELEFAX: (248) 5395055

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:
LENGTH: 1782 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-604-978-2

Query Match 6.9%; Score 121; DB 4; Length 1782;

Best Local Similarity 83.4%; Pred. No. 4e-19;
Matches 176; Conservative 0; Mismatches 25; Indels 10; Gaps 3;

QY 1 TTGTTTTTAAACAACATGTTTATTAGAAAAGT-AAAAATATTGCATAGGTCTTAGTACTT 59
Db 1767 TTGTTTTTAAACAACATGTTTATTAGAAAAGTAAAAAATATTGCATAGTCTTAACTT 1708
QY 60 GAACATCAAGTGATTTATGAAACCGTGAG-----TATCTTCATGTAACAGTTCTAGATG 114
Db 1707 GAACATCAAGTGATTTATGAAACAGTGAGTATCTTATCTTCATGTAACAGTTCTAGATG 1648
QY 115 GAAGACCCAGGTGGCGCTCTCTGCGGGAGAGGGTTCCAGCCCCCAGCCCCCTAGCCC 174
Db 1647 GAAGACCCAGATGGCACTCTCTCCCGGGAGGGGTTCCAGCCCCCAGCCCCCTCTC-----AGCC 1592

QY 175 CATCCCTCACAGCTCACTCTCTCCAGTACAC 205
Db 1591 CTTCCCTGCCAGCTCAACTCTGCAGTACAC 1561

RESULT 14

US-09-604-728-2/c
; Sequence 2, Application US/09604728
; Patent No. 6555667
; GENERAL INFORMATION:
; APPLICANT: Einat, Paz
; Skaliter, Rami
; TITLE OF INVENTION: HYPOXIA-REGULATED GENES
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: KOHN & ASSOCIATES
; STREET: 30500 No. 6555667thwestern Hwy., Suite 401
; CITY: Farmington Hills
; STATE: Michigan
; COUNTRY: U.S.
; ZIP: 48334

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/604,728
FILING DATE: 28-Jun-2000
CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/09/138,112
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Kohn, Kenneth I.
REGISTRATION NUMBER: 30,955
REFERENCE/DOCKET NUMBER: 0168.00034
TELECOMMUNICATION INFORMATION:
TELEPHONE: (248) 539-5050
TELEFAX: (248) 5395055

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:
LENGTH: 1782 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-604-728-2

Query Match 6.9%; Score 121; DB 4; Length 1782;

Best Local Similarity 83.4%; Pred. No. 4e-19;
Matches 176; Conservative 0; Mismatches 25; Indels 10; Gaps 3;

QY 1 TTGTTTTTAAACAACATGTTTATTAGAAAAGT-AAAAATATTGCATAGGTCTTAGTACTT 59
Db 1767 TTGTTTTTAAACAACATGTTTATTAGAAAAGTAAAAAATATTGCATAGGTTCTTAATACTT 1708
QY 60 GAACATCAAGTGATTTATGAAACCGTGAG-----TATCTTCATGTAACAGTTCTAGATG 114
Db 1707 GAACATCAAGTGATTTATGAAACAGTGAGTATCTTATCTTCATGTAACAGTTCTAGATG 1648
QY 115 GAAGACCCAGGTGGCGCTCTCTGCGGGAGAGGGTTCCAGCCCCCAGCCCCCTAGCCC 174
Db 1647 GAAGACCCAGATGGCACTCTCTCCCGGGAGGGGTTCCAGCCCCCAGCCCCCTCTC-----AGCC 1592
QY 175 CATCCCTCACAGCTCACTCTCTCCAGTACAC 205
Db 1591 CTTCCCTGCCAGCTCAACTCTGCAGTACAC 1561

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 26, 2003, 13:17:02 ; Search time 3862 Seconds
(without alignments)
11025.749 Million cell updates/sec

Title: US-10-027-859-1

Perfect score: 1752
Sequence: 1 ttgttttaacaacatgtt.....aaaaaaaaaaaaaaaaaaaaa 1752

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST.*

1: em_estba.*
2: em_esthum.*
3: em_estin.*
4: em_estmu.*
5: em_estov.*
6: em_estpl.*
7: em_estro.*
8: em_hcc.*
9: gb_est1.*
10: gb_est2.*
11: gb_hcc.*
12: gb_est3.*
13: gb_est4.*
14: gb_est5.*
15: em_estfun.*
16: em_estom.*
17: em_gss_hum.*
18: em_gss_inv.*
19: em_gss_pln.*
20: em_gss_vrt.*
21: em_gss_fun.*
22: em_gss_man.*
23: em_gss_mus.*
24: em_gss_pro.*
25: em_gss_rtd.*
26: em_gss_phg.*
27: em_gss_vrt.*
28: gb_gss1.*
29: gb_gss2.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
c 1	472.4	27.0	599	12	BI438572 iC24a08.x
c 2	452.2	25.8	590	14	CA867675 i31f10.y
c 3	415.8	23.7	495	14	CA771025 i069g10.x
c 4	415.4	23.7	584	14	CA771327 i069g10.y

c 5	414.4	23.7	472	14	CA867392	CA867392 ir31f10.x
c 6	404	23.1	436	9	AM488372	AM488372 UI-M-BH3-
c 7	373.2	21.3	516	12	BI194956	BI194956 602943780
c 8	361.6	20.6	531	10	BE948739	BE948739 UI-M-BH3-
c 9	358.2	20.4	379	10	BE948133	BE948133 UI-M-BH3-
c 10	356.4	20.3	632	10	BE262829	BE262829 601151586
c 11	354	20.2	1120	13	BU603585	BU603585 AGENCOURT
c 12	334.6	19.0	649	10	BE312349	BE312349 601148529
c 13	333.8	19.1	451	14	CB069127	CB069127 i813c07.x
c 14	329	18.8	582	14	CB069363	CB069363 i813c07.y
c 15	318	18.2	458	10	BF470469	BF470469 UI-M-BH3-
c 16	288.6	16.5	624	13	BQ777436	BQ777436 i148b07.y
c 17	283.4	16.2	635	28	AQ937962	AQ937962 NB6-699R
c 18	271.2	15.5	480	13	BQ788110	BQ788110 i148b07.x
c 19	270	15.4	321	12	BM128563	BM128563 i12h07.y
c 20	260.8	14.9	430	12	BI963862	BI963862 ie65f04.x
c 21	255.8	14.6	675	12	BM440515	BM440515 pgrln.pk0
c 22	251.2	14.3	377	28	AQ939613	AQ939613 NR3-088R
c 23	244.2	13.9	671	13	BU417183	BU417183 603671376
c 24	242.6	13.8	289	12	BM128343	BM128343 i12h07.x
c 25	233.4	13.3	322	9	AI577542	AI577542 UI-R-Y0-v
c 26	233.4	13.3	675	13	BQ199014	BQ199014 UI-R-EA0-
c 27	233.4	13.3	700	13	BQ209348	BQ209348 UI-R-DY1-
c 28	233.4	13.3	759	14	CA509882	CA509882 UI-R-F50-
c 29	231.8	13.2	346	9	AI555662	AI555662 UI-R-C2p-
c 30	231.8	13.2	499	9	AI144692	AI144692 UI-R-BT0-
c 31	231.8	13.2	656	14	CD371354	CD371354 UI-R-DY0-
c 32	231.8	13.2	673	12	BM384149	BM384149 UI-R-D20-
c 33	230.2	13.1	290	9	AI555747	AI555747 UI-R-C2p-
c 34	225.4	12.9	358	9	AI235276	AI235276 EST331838
c 35	225.4	12.9	456	9	AI408716	AI408716 EST337007
c 36	220.8	12.6	694	13	BQ210597	BQ210597 UI-R-DY1-
c 37	219.8	12.5	634	28	AQ939612	AQ939612 NR3-088C
c 38	216.8	12.4	251	9	AI105175	AI105175 EST214464
c 39	205	11.7	524	9	AI794504	AI794504 fc45b11.y
c 40	203	11.6	1866	11	AK050922	AK050922 Mus muscu
c 41	203	11.6	3064	11	AK078113	AK078113 Mus muscu
c 42	198.8	11.3	466	13	BQ270970	BQ270970 i11f04.x
c 43	198.2	11.3	602	14	CB582161	CB582161 AMGNNUC:N
c 44	198.2	11.3	616	14	CB579500	CB579500 AMGNNUC:N
c 45	197.8	11.3	709	14	CB269538	CB269538 108445 H

ALIGNMENTS

RESULT 1
BI438572/c BI438572 599 bp mRNA linear EST 21-AUG-2001
LOCUS iC24a08.x3 HR85 islet Homo sapiens cDNA 3', similar to TR:070132
DEFINITION O70132 ETS DOMAIN TRANSCRIPTION FACTOR PET-1. ; mRNA sequence.
ACCESSION BI438572 GI:15263262
VERSION EST.
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 599)
Morton, D., Brown, J., Kenty, G., Permutt, A., Lee, C., Kaestner, K.,
Lemishka, I., Searce, M., Brestelli, J., Gradwohl, G., Clifton, S.,
Haller, R., Marra, M., Pape, D., Wylie, T., Martin, J., Blistain, A.,
Schmitt, A., Theising, B., Ritter, E., Ronko, I., Bennett, J., Cardenas
M., Gibbons, M., McCann, R., Cole, R., Tsagareishvili, R., Williams, T.,
Jackson, Y. and Bowers, Y.
TITLE Endocrine Pancreas Consortium
JOURNAL Unpublished
COMMENT Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue
Endocrine Pancreas Consortium
Harvard University, Howard Hughes Medical Institute
Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,
MA 02138
Tel: 617-495-1812

Fax: 617-495-8557
 Email: dmelton@biohp.harvard.edu
 Library was constructed by Dr. Hiroshi Inoue DNA sequencing by:
 Washington University Genome Sequencing Center For information on
 obtaining a clone please contact: Dr. Hiroshi Inoue
 (hinoue@im.wustl.edu)
 High quality sequence stop: 475.
 Location/Qualifiers
 1..599
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /tissue_type="Purified pancreatic islet"
 /lab_host="DH10B"
 /clone_lib="HR85 islet"
 /note="Organ: Pancreas; Vector: pBluescript SK(-); Site_1:
 NotI; Site_2: XhoI; cDNA made by oligo-dT priming.
 Size selected on agarose gel. Average insert size -1kb. 5'
 XhoI site was destroyed after directional cloning.
 Amplified once. Contact information: Hiroshi Inoue, MD,
 Metabolism Div. (Alan Permutt Lab), Washington University
 School of Medicine, Box 8127, 660 South Euclid Ave., St.
 Louis, MO 63110, E-mail: hinoue@imgate.wustl.edu, Tel:
 314-362-1916, Fax: 314-747-2692."
 82 a 191 c 223 g 100 t 3 others

FEATURES source

Query Match 27.0%; Score 472.4; DB 12; Length 599;
 Best Local Similarity 88.8%; Pred. No. 3e-39;
 Matches 532; Conservative 0; Mismatches 64; Indels 3; Gaps 2;
 QY 289 CCGGACCCGGCGCTTACTTCTCCCTGTACCCACAGTACACCGCGGGTATCGGC 348
 Db |||||
 QY 599 CCGGACCCGGCGCGTTTCTTCTCTGTGTACCCCGGTGCGTGGCGGGGATCGGT 540
 Db |||||
 QY 349 ACC-CAAAGCGCAAAGTGAAGTGTGCGCCCGCCCT--GCGGTCCCCCCCCATCTCCCAACGCC 405
 Db |||||
 QY 539 GCCACCGAGCGAAAGCGTCTGCGCCCGCTTCCCGGTGTCGCCCATCTCCCAACGCC 480
 QY 406 CAGTCCCGGACGATGAGACAGAGCGGACCTCCAGCCCTCTGATCAACATGTAC 465
 Db |||||
 QY 479 CAGTCCCGGCGGATGAGACAGAGCGGCGCTTCCAGCCCTCTGATCAACATGTAC 420
 QY 466 TTACCAGATCCCGTGGAGATGGTCTTTTAAAGGAGGAGAGCGCCGAGCTGGGGCGG 525
 Db |||||
 QY 419 CTGCCAGATCCGTCGGAGACGGTCTTCAAGAGCGGAGAGAACCCGAGCTGGGGCGG 360
 QY 526 CTGAGCCCTGGGTACAGAAAGGCGAGCGGAGATCCAGTTGTGGCAGTTTCTACTGGAG 585
 Db |||||
 QY 359 CTGAGCCCGCGGTTTCAAGAAAGGCGAGCGAGATCCAGTGTGGCAGTTTCTGCTGGAG 300
 QY 586 CTGCTGGCAGACCGCGGAAACCGCGTGCATCGGTGGAGGGGCGGCGCAGCGAGTTC 645
 Db |||||
 QY 299 CTGCTGGTGTACCGCGGAAACCGCGTGCATCGGTGGAGGGGCGGTACCGGCGAGTTC 240
 QY 646 AAGCTCACCGACCCCGACGAGTGGCGGACGCTGGGGCGAGCGCAAGAGCAAGCCCAAT 705
 Db |||||
 QY 239 AAGCTCACCGACCCCGACGAGTGGCGGCGGTGGGGCGAGCGCAAGAGCAAGCCCAAC 180
 QY 706 ATGAACACTACGAAAGCTAAGTTCGAGCATGCGGTACTACTACGACAAACATCATGAGC 765
 Db |||||
 QY 179 ATGAACACTACGAAAGCTAAGTTCGAGCATGCGGTACTACTACGACAAACATCATGAGC 120
 QY 766 AAGTGGACGCAAGCGCTACGCTTACCGCTTTGACTTCCAGGGCTTGGCAGAGCTTGC 825
 Db |||||
 QY 119 AAGTGGATGCAAGCGCTACGCTTACCGCTTCCAGGGCTTGGCAGAGCTTGC 60
 QY 826 CAGCCACCAACCGGCGACGCCCAAGCGCCCGCTGCGCGCGCGGAGCGGACGCGCGC 884
 Db |||||
 QY 59 CAGCGCGCGCGCGGACGCTCATGCGCGCGCGCGAGCTGTGCGCGCGCGCGCGCGC 1

RESULT 2

CA867675 590 bp mRNA linear EST 20-DEC-2002
 LOCUS ir3if10.y1 HR85 islet Homo sapiens cDNA clone IMAGE:6546788 5'
 DEFINITION similar to TR:070132 070132 ETS DOMAIN TRANSCRIPTION FACTOR PET-1.
 ; mRNA sequence.
 ACCESSION CA867675
 VERSION CA867675.1 GI:27319224
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 REFERENCE 1 (bases 1 to 590)
 AUTHORS Melton,D., Brown,J., Kenty,G., Permutt,A., Lee,C., Kaestner,K.,
 Lemishka,L., Searce,M., Brestelli,J., Gradwohl,G., Clifton,S.,
 Hillier,L., Marra,M., Pape,D., Wyllie,T., Martin,J., Blisstein,A.,
 Schmitt,A., Theising,B., Ritter,E., Ronko,I., Bennett,J., Cardenas,
 M., Gibbons M., McCann,R., Cole,R., Tsagarishvili,R., Williams,T.,
 Jackson,Y. and Bowers,Y.
 TITLE Endocrine Pancreas Consortium
 JOURNAL Unpublished
 COMMENT Other ESTs: ir3if10.x1
 Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue
 Endocrine Pancreas Consortium
 Harvard University, Howard Hughes Medical Institute
 Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,
 MA 02138
 Tel: 617-495-1812
 Fax: 617-495-8557
 Email: dmelton@biohp.harvard.edu
 Library was constructed by Dr. Hiroshi Inoue DNA sequencing by:
 Washington University Genome Sequencing Center For information on
 obtaining a clone please contact: Dr. Hiroshi Inoue
 (hinoue@im.wustl.edu)
 Seq primer: -40UP from Gibco
 High quality sequence stop: 455.
 Location/Qualifiers
 1..590
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:6546788"
 /tissue_type="Purified pancreatic islet"
 /lab_host="DH10B"
 /clone_lib="HR85 islet"
 /note="Organ: Pancreas; Vector: pBluescript SK(-); Site_1:
 NotI; Site_2: XhoI; cDNA made by oligo-dT priming.
 Size selected on agarose gel. Average insert size -1kb. 5'
 XhoI site was destroyed after directional cloning.
 Amplified once. Contact information: Hiroshi Inoue, MD,
 Metabolism Div. (Alan Permutt Lab), Washington University
 School of Medicine, Box 8127, 660 South Euclid Ave., St.
 Louis, MO 63110, E-mail: hinoue@imgate.wustl.edu, Tel:
 314-362-1916, Fax: 314-747-2692."
 103 a 224 c 179 g 84 t
 BASE COUNT
 ORIGIN
 Query Match 25.8%; Score 452.2; DB 14; Length 590;
 Best Local Similarity 89.5%; Pred. No. 3.4e-37;
 Matches 509; Conservative 0; Mismatches 58; Indels 2; Gaps 2;
 QY 240 CCCCCTCCCTCTCACACACCCCTGGCTGCCGCTCCCGCCGAGGCTTCCCGACCCCGC 299
 Db |||||
 QY 22 CCCCCGAAACCCCGCACCGCCGGGTGCCCCCTCCCGCAGGCCCCCTGCGGACCCGCGC 81
 QY 300 GGCGTCTACTTCTCCCTGTACCCACAGTACACCGGCGGGGTATCGGCACC-CAAAGCG 358
 Db |||||
 QY 82 GGCGTCTTCTCTCTGTGTACCCCGCGTCCGTTGGGGGGGGATCGGTGCACCGACGCG 141
 QY 359 CAAAG-CTGAGTGTCCCCCGCGTGGGTGCCCCCATCTCCACCGCCAGTCCCCGGCA 417
 Db |||||
 QY 142 CAAAGCTGCTCGCCCCCTTCCCGTCCCGCCCATCTCCACCGCCAGTCCCCCGCGC 201

352	TTCAAGGACGGGAAGAACCCGAGCTGGGGCCGCTGAGCCCGCGGTTTCAGAAAGGCAGC	293
553	GGGCGATCCAGATTGTGGCAGTTTCTACTGGAGCTGCTGGCGAGACC GCGGAGCCGCGC	612
292	GGACAGATCCAGCTGTGGCGAGTTTCTGCTGGAGCTGCTGGCTGACCCGCGGAGCCGCGC	233
613	TGATCCGCTGGAGGGCGGCCACGGCGAGTTTCAAGCTCACCGGACCCGCGAGGTCGGC	672
232	TGCATCCGCTGGGAGGGCGGTCACGGCGAGTTTCAAGCTCACGGACCCGCGAGGTCGGC	173
673	CGAGCTGGGGCGAGCGCAAGAGCAAGCCCAATATGACTACGACAGCTAAGTCGAGCA	732
172	CGCGGTGGGGCGAGCGCAAGAGCAAGCCCAATGAATACGACAGCTGAGCGCGCGCC	113
733	CTCGCTTACTACTACGACAAAACATCATGAGCAAGGTGACGCGAAGCGCTACGCGCTAC	792
112	CTCGCTTACTACTACGACAAAGCATCATGAGCAAGGTGCATGGCAAGCGCTACGCGCTAC	53
793	CGCTTTGACTTCAGGGCGCTGGCACAGGCTTGCCAGGCCACACCCCGCGCAGC	844
52	CGCTTTCGACTTCAGGGCGCTGGGGCAGGCTGCGAGCGCGCGCGCGCGCAGC	1

RESULT 6					
AW488372/c					
LOCUS	AW488372	436 bp	mRNA	linear	EST 24-FEB-2000
DEFINITION	UI-M-BH3-arp-g-06-0-UI.s1 NIH BMAP_M_S4 Mus musculus cDNA clone UI-M-BH3-arp-g-06-0-UI 3', mRNA sequence.				

FEATURES
source

Location/Qualifiers

1. 436

/organism="Mus musculus"

/mol_type="mRNA"

/strain="C57BL/6J"

/db_xref="taxon:10090"

/clone="UI-M-BH3-arp-g-06-0-UI"

/dev_stage="27-32 days"

/lab_host="DH10B (Life Technologies)"

/clone_lib="NIH_BMAP_M_S4"

/note="Vector: pT7T3D-Pac (Pharmacia) with a modified polylinker. Site 1: Not I; Site 2: Eco RI; The NIH_BMAP_M_S4 library is a subtracted library of a series, ultimately derived from a mixture of individually tagged

normalized libraries from ten regions of the mouse brain (cerebellum, brain stems, olfactory bulbs, hypothalamus, cortex, amygdala, basal ganglia, pineal gland, striatum, hippocampus) after a series of subtractions to reduce the representation of cDNAs from which ESTs had already been generated. The following serially subtracted libraries were generated in this process: NIH_BMAP_M_S4, NIH_BMAP_M_S3.3, NIH_BMAP_M_S3.2, NIH_BMAP_M_S3.1, NIH_BMAP_M_S2, NIH_BMAP_M_S1. The subtracted library (NIH_BMAP_M_S4) was constructed as follows: PCR amplified cDNA inserts from NIH_BMAP_M_S3.3, NIH_BMAP_M_S3.2, and NIH_BMAP_M_S3.1 clones from which 3' ESTs had been derived was used as a driver in a hybridization with a pool of the NIH_BMAP_M_S3.3, NIH_BMAP_M_S3.2, and NIH_BMAP_M_S3.1 libraries in the form of single-stranded cDNAs. The remaining single-stranded circles (subtracted library) was purified by hydroxyapatite column chromatography, converted to double-stranded circles and electroporated into DH10B bacteria (Life Technologies) to generate the NIH_BMAP_M_S4 library. This procedure has been previously described (Bonaldo, Lennon and Soares, Genome Research 6:791-805 1996).

BASE COUNT	63 a	142 c	141 g	90 t
ORIGIN				
Query Match	23.1%	Score 404;	DB 9;	Length 436;
Best Local Similarity	95.4%	Pred. No. 3.4e-32;		
Matches 416;	Conservative 0;	Mismatches 20;	Indels 0;	Gaps 0;
QY	427	CAGAGCGGCACCTCC	CAGGCCCTGCTGATCAACATGTACTAC	CAGATCCCGTCGGAGAT 486
Db	436	CGCAGCAGGGCTCC	CAGGCCCTGCTGATCAACATGTACTCC	CAGATCCCGTCGGAGAT 377
QY	487	GGTCTTTTTAAGGAAGG	AAGCCCGAGCTGGGGCCCGCTGAGCCCTCG	CGGTACAGAAA 546
Db	376	GGTCTTTTTAAGGAAGG	AAGCCCGAGCTGGGGCCCGCTGAGCCCGCGGT	ACAGAAA 317
QY	547	GGCAGGGGCAGATC	AGTTGTGGAGTTTCTACTGGAGCTCTGGCAG	ACCCCGCGAAC 606
Db	316	GGCAGGGGCAGATC	AGTTGTGGAGTTTCTCTGGAGCTCTGGCAG	ATCGCGGAAC 257
QY	607	GCGGCTGCATCG	TGGAGGGCGGCACCGCGAGTTCAAGCTCAC	CGACCCCGACGAG 666
Db	256	GCGGTTCATCG	TGGAGGGCGGCACCGCGAGTTCAAGCTCAC	CGACCCCGACGAG 197
QY	667	GTGGCGCGACGCT	TGGGGGAGCGCAAGAGCCCAATATCAACTAC	GACAAGCTTAAGT 726
Db	196	GTGGCGCGCGCT	TGGGGGAGCGCAAGAGCCCAATATCAACTAC	GACAAGCTTAAGT 137
QY	727	CGAGCACTGCGCT	TACTACTACGACAAAACATCATGAGCAAGGTGC	CGGCAAGCGCTAC 786
Db	136	CGCGCGCTGCGCT	ACTACTACGACAAAACATCATGAGCAAGGTGC	CGGCAAGCGCTAC 77
QY	787	GCCTACCGCTTTG	ACTTCCAGGGCTTGGCAGAGCTTGCAGCCAC	CACCACCCCGCACGCC 846
Db	76	GCCTACCGCTTTG	ACTTCCAGGGCTTGGCAGAGCTTGCAGCCAC	CACCACCCCGCACGCC 17
QY	847	CACCGCGCGCTG	CCG 862	
Db	16	CACCGCGCTG	CGGCGG 1	

RESULT 7	BI194956	516 bp	mRNA	linear	EST 10-JUL-2001
LOCUS	BI194956				
DEFINITION	602943780F1 NIH_MGC_19 Homo sapiens cDNA clone IMAGE:5092074 5', mRNA sequence.				
ACCESSION	BI194956				
VERSION	BI194956.1				GI:14649976
KEYWORDS	EST				


```
BASE COUNT      TAG_SEQ=None found"
ORIGIN          85 a 172 c 170 g 103 t 1 others

Query Match      20.6%; Score 361.6; DB 10; Length 531;
Best Local Similarity 94.8%; Pred. No. 5.6e-28;
Matches 384; Conservative 0; Mismatches 28; Indels 1; Gaps 1;

QY 459 CATGTACTACAGATCCCGTGGAGATGGTCTTTTAAAGGAAGGAAGAGCCGAGCTG 518
DB 405 CTTNTGCGCGCAGATCCCGTGGAGATGGTCTTTTAAAGGAAGGAAGAGCCGAGCTG 346
QY 519 GGGCGCGTGGCTGCGGTACAGAAAGGAGCGGGGAGATCCAGTTGTGCGAGTTTCT 578
DB 345 GGGCGCGTGGCTGCGGTACAGAAAGGAGCGGGGAGATCCAGTTGTGCGAGTTTCT 286
QY 579 ACTGAGCTGTGGCAGACCGCGCAACCGCGCTGCATCGTGGAGGGCGGCACGG 638
DB 285 CTTGAGCTGTGGCAGATCGCGCAACCGCGTGTGCATCGTGGAGGGCGGCACGG 226
QY 639 CGAGTTCAAGCTACCGACCGCGAGCGAGTGGCGGCGAGCTGGGCGGAGCGCAAGACAA 698
DB 225 CGAGTTCAAGCTACCGACCGCGAGCGAGTGGCGGCGAGCGCGAGCGCAAGACAA 166
QY 699 GCCCATATGACTACGACAGCTTAAGTCGAGCACTGCGCTACTACTACGACAAAAACAT 758
DB 165 GCCTAATGACTACGACAGCTTAAGTCGCGCGTGGCGCTACTACTACGACAAAAACAT 106
QY 759 CATGAGCAAGTGCACGCGCAAGCGCTACCGCTTACCGCTTTGACTT-CCAGGCGCTGGCAC 817
DB 105 CATGAGCAAGTGCACGCGCAAGCGCTACCGCTTACCGCTTTGACTTCCAGGCGCTGGCGC 46
QY 818 AGGCTTGCAGCGACCGACCGCGGAGCGCGCGCGCGCGCGCGCGCGCG 862
DB 45 AGGCTTGCAGCGACCGACCGCGGAGCGCGCGCGCGCGCGCGCGCG 1

RESULT 9
BE948133/c
LOCUS
DEFINITION
  UI-M-BH3-awn-c-08-0-UI-s1 NIH BMAP M S4 Mus musculus cDNA clone
  UI-M-BH3-awn-c-08-0-UI 3', mRNA sequence.
ACCESSION
  BE948133.1 GI:10525892
VERSION
  EST.
KEYWORDS
  Mus musculus (house mouse)
SOURCE
  Mus musculus
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
  (Bases 1 to 379)
  Bonaldo, M.F., Lennon, G. and Soares, M.B.
  Normalization and subtraction: two approaches to facilitate gene
  discovery
  Genome Res. 6 (9), 791-806 (1996)
  97044477
  8889548
  Contact: Chin, H
  National Institute of Mental Health
  6001 Executive Blvd. Room 7N-7190, MSC 9643, Bethesda, MD
  20892-9643, USA
  Tel: 301 443 1706
  Fax: 301 443 9890
  Email: MEST@mail.nih.gov
  Oligo-dT track not found, Not 1 site shown in beginning of sequence
  is likely internal to the message. cDNA Library Preparation: M.B.
  Soares Lab Clone distribution: Researchers may obtain BMAP cDNA
  clones from RESEARCH GENETICS. It should be noted that Bento Soares
  is generating a small number of additional specialized
  non-redundant arrays of BMAP cDNAs whose availability will be
  considered under appropriate and limited collaborative arrangements
  Seq primer: M13 Forward
  POLYA=No.
  Location/Qualifiers

FEATURES
  BASE COUNT      54 a 128 c 118 g 79 t
  ORIGIN
  Query Match      20.4%; Score 358.2; DB 10; Length 379;
  Best Local Similarity 96.6%; Pred. No. 1.7e-27;
  Matches 366; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

  QY 484 GATGCTCTTTTAAAGGAAGGAAGAGCGCGAGCTGGGGCGCGCTCAGCCCTGCGGTACAG 543
  DB 379 GATGCTCTTTTAAAGGAAGGAAGAGCGCGAGCTGGGGCGCGCTCAGCCCTGCGGTACAG 320
  QY 544 AAAGCGAGCGGCGAGATCCAGTTGTGCGAGTTTCTACTGGAGCTGCTGGCAGACCGCGCG 603
  DB 319 AAAGCGAGCGGCGAGATCCAGTTGTGCGAGTTTCTCTCTGGAGCTGCTGGCAGATCGCGCG 260
  QY 604 AACGCCGCTGCATCGCTGGAGGGCGGCACCGCGAGTTCAAGCTCACCAGCCCGGAC 663
  DB 259 AACGCCGCTGCATCGCTGGAGGGCGGCACCGCGAGTTCAAGCTCACCAGCCCGGAC 200
  QY 664 GAGTGGCGCGACGCTGGCGGCGAGCGCAAGAGCGCCCATATGAACTACGCAAGCTA 723
  DB 199 GAGTGGCGGCGCGCTGGCGGCGAGCGCAAGAGCGCCCTAACATGAACTACGCAAGCTA 140
  QY 724 AGTCGACACTGCGCTACTACTACGACAAAAACATATGAGCAAGGTGCACGGCAAGCGC 783
  DB 139 AGTCGCGCGCTGCGCTACTACTACGACAAAAACATATGAGCAAGGTGCACGGCAAGCGC 80
  QY 784 TAGCGCTACCGCTTTGACTTCCAGGGCTGGCAGAGCTTGCAGCCACCAACCCCGCGCAC 843
  DB 79 TAGCGCTACCGCTTTGACTTCCAGGGCTGGCAGAGCTTGCAGAGCTTGCAGCCACCAACCCCGCGCAC 20
  QY 844 GCCACGCGCGCGCTGCGCG 862
  DB 19 GCCACGCGCGCTGCGCGCG 1

RESULT 10
```



```

BE262829
LOCUS BE262829 632 bp mRNA linear EST 26-OCT-2000
DEFINITION G0115186F1 NIH_MGC_19 Homo sapiens cDNA clone IMAGE:3507620 5',
mRNA sequence.
ACCESSION BE262829
VERSION BE262829.1 GI:9136188
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE NIH-MGC http://mgc.nci.nih.gov/.
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Ling Hong/Rubin Laboratory
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
Plate: LCM187 row: 1 column: 21
High quality sequence start: 27
High quality sequence stop: 616.
Location/Qualifiers
1..632
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:3507620"
/tissue_type="neuroblastoma"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_19"
/note="Organ: brain; Vector: pOTB7; Site 1: XhoI; Site 2:
EcoRI; cDNA made by oligo-dT priming. Directionally
cloned into EcoRI/XhoI sites using the following 5'
adaptor: GGCACGAG(G). Library constructed by Ling Hong
in the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies)."
Note: this is a NIH MGC library."
BASE COUNT 90 a 246 c 203 g 93 t
ORIGIN

Query Match 20.3%; Score 356.4; DB 10; Length 632;
Best Local Similarity 83.1%; Pred. No. 1.6e-27;
Matches 505; Conservative 0; Mismatches 86; Indels 17; Gaps 8;

QY 515 GCTGGGGCGGCTGAGCCCTGCGGTACAGAAAGGCGGCGAGATCCAGTTGTGGCAGT 574
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
29 GCGGAGGGCGCTGAGCCCGCGGTTTCAGAA----AGCGACAGATCCAGCTGTGGCAGT 84
QY 575 TTCTACTGAGCTGTGGCAGACCCGCGAAACCGCGGTGTCATCGGTGGGAGGGCGGCC 634
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
85 TTCTGCTGAGCTGTGCTGACCGCGGCGAAGCCCGGCTGTCATCGGTGGGAGGGCGGTC 144
QY 635 ACGGCGAGTTCAAGCTCACCGACCCCGAGAGGTGGCGGACGCTGGGCG-GAGCGCAAG 693
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
145 ACGGCGAGTTCAAGCTCAAGGACCCCGAGGTGGCGGCGGTGGGCGGTGAGCGCAAG 204
QY 694 AGCAAGCCCAATATGAACTACGACAAAGCTTAAGTCGAGCACTGCGCTACTACTACGACAA 753
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
205 AGCAAGCCCAATGAACTACGACAAAGCTGAGCGCGCGCTGCGCTACTACTACGACAA 264
QY 754 AACATCATGAGCAAGGTGACGCGCAAGCGCTACCGCTTTGACTTCCAGGCGCTG 813
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
265 AACATCATGAGCAAGGTGATGCAAGCGCTACCGCTACCGCTTTCAGTTCAGGCGCTG 324
QY 814 GCACAG-GCTTGCACGACCAACCCGCGGACCGCCACGCGC-GCCGCTGCCGCGCGGCGAG 871
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
325 GCGCAGTGCTGCGAGCCCGCGCGCGCACGCTATGCTCTGCCGCGCAGCTGTGTCGCG 384

```

```

QY 872 CGGCGAGC-CGCCGCCAGGATGGCGCAGCTTTTAAAGCT-CCCGGTGGTCTGGCTCCACT 929
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
385 CGCGCGCTGGCGCGCCAGGACGGCGGCTCTACAAGCTCGCGCCTGGCCTGCCCCGCT 444
QY 930 GCCCTTCCCGCGCCTCTCCAACTCAACCTTATGCGAGCCTGGCGCGGCTGGCGGCC-G 988
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
445 GCCCTTCCCGCGCCTCTCCAACTCAACCTCAATGCGCGCCTCGCGCGGCTGGCGCCTG 504
QY 989 CTGGCTTCTTTACTGGCTGTGTCACAAAGCCACCGCCGCTCGCGCGGCACAC----- 1041
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
505 CGCGCTTCTCTACTGGCTGTGCGCGCGGTGGCGCTGCGCTGACCGCTGCGCGCGCACCT 564
QY 1042 GCTGGCTTACCAACCCCGCGCTTGCAGCCCTTCCCGCGCCTTTTTCGCGCGGTGGCC 1101
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
565 GCGCGCTCTACCCAGTTCAGCTTGCAGCCCGCGCGCGCTTTCGCGCGCTGTGGCC 624
QY 1102 GCGGCTTC 1109
DB 625 GCAGCTC 632

BU603585 1120 bp mRNA linear EST 20-SEP-2002
AGENCOURT_8934291 NIH_MGC_142 Homo sapiens cDNA clone IMAGE:6499070
5', mRNA sequence.
ACCESSION BU603585
VERSION BU603585.1 GI:23255344
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 1120)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: NCI
cDNA Library Preparation: Michael Brownstein Laboratory
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LCM2682 row: 1 column: 15
High quality sequence stop: 290.
Location/Qualifiers
1..1120
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6499070"
/tissue_type="mixed (pool of 40 RNAs)"
/lab_host="DH10B (T1-phage-resistant)"
/clone_lib="NIH_MGC_142"
/note="Vector: pDNR-LIB; Site 1: SfiI (ggccattatggcc);
Site 2: SfiI (ggcgccctcgcc); Double-stranded cDNA was
prepared from a pool of 40 cell line polyA+ RNAs (bladder
2%, blood - 33.4%, brain - 5.6%, breast - 12.5%, colon -
4%, connective tissue - 1.4%, eye - 1%, intestine - 2.6%,
kidney - 2.2%, liver - 5.7%, lung - 10.8%, NK-cell - 5.2%,
ovary - 4%, pharynx - 2.5%, prostate - 4.3%, salivary
gland - 1.3%, and skin - 2.3%). 5' and 3' adaptors were
used in cloning as follows:
5'-RAGCAGTGTATCAGCAGAGTGGCCATTACGGCGGG-3' and
5'-ATCTAGAGCGGAGCGCGCATG-dT(30)NN-3'. Full-length
enriched library was constructed using the Clontech
Creator SMART kit and size-selected to contain the >0.5 kb
size fraction (other fractions present in NIH MGC 141).
Library created in the laboratory of M. Brownstein (NIH),

```


BASE COUNT 174 a 397 c 353 g 193 t 3 others
 ORIGIN NIH). Note: this is a NIH MGC Library."

Query Match 20.2%; Score 354; DB 13; Length 1120;
 Best Local Similarity 87.6%; Pred. No. 1.7e-27; Mismatches 50; Indels 10; Gaps 3;
 Matches 424; Conservative 0; Mismatches 50; Indels 10; Gaps 3;
 QY 428 AGAGCGCGACCTCCACAGCCCTGCTGATCAATGATACCTACAGATCCCGTCGGAGATG 487
 DB 3 AGAGCGCGCCCTCCACAGCCCTGCTGATCAATGATACCTGCGAGATCCCGTCGGAGCG 62
 QY 488 GTCTTTTAAAGAGGAAGAGCCAGCTGGGGCGCTGAGCCCTGCGGTACAGAAAG 547
 DB 63 GTCTCTTCAAGGACGGAAGAACCCGAGCTGGGGCGCTGAGCCCGCGGTTTCAAGAA 122
 QY 548 GCAGCGGGCAGATCCAGTTGTGGCAGTTTCTACTGAGCTGTGGCAGACCGCGGAACG 607
 DB 123 GCAGCGGACAGATCCAGCTGTGGCAGTTTCTGCTGAGCTGTGGCTGACCGCGGAACG 182
 QY 608 CCGGCTGTCATCGCTGGGAGGGCGCCACGCGAGTTCAAGCTCACGACCCCGAGAGG 667
 DB 183 CCGGCTGTCATCGCTGGGAGGGCGGTACGCGAGTTCAAGCTCACGACCCCGAGAGG 242
 QY 668 TGGCGGAGCTGGGGCGAGCGCAAGCAAGCCCAATATGAATACGACAAGCTTAAGTC 727
 DB 243 TGGCGGCGGTGGGGCGAGCGCAAGCAAGCCCAATATGAATACGACAAGCTTAAGTC 302
 QY 728 GAGCACTGCGCTACTACTAGCAAA-AAACATCATGAGCAAGGTGACGCAAGCGCTAC 786
 DB 303 GCGCCCTGCGCTACTACTAGCAAA-AAACATCATGAGCAAGGTGACGCAAGCGCTAC 362
 QY 787 GCTACCGCTTTGACTTCCA-GGGCTGGCACAGGCTTCCAGCCACCCGCGCACGC 845
 DB 363 GCTACCGCTTTCAGTTCAGGGGCTTGGCGAGGCTGCCAGCGCGCGCGCACGT 422
 QY 846 CCAGCGCGCGTGGCGCGCGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 905
 DB 423 TCATGCCCCCG-----GCGCGAGCTGGTGGCGCGCGCGCGCGCGCGCGCGCGCG 474
 QY 906 GCTC 909
 DB 475 GCGC 478

RESULT 12
 BE312349 649 bp mRNA linear EST 26-OCT-2000
 LOCUS 601148529F1 NIH_MGC_19 Homo sapiens cDNA clone IMAGE:3163855 5',
 DEFINITION mRNA sequence.
 ACCESSION BE312349
 VERSION BE312349.1 GI:91130901
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 NIH-MGC http://img.nci.nih.gov/.
 1 (bases 1 to 649)
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgaabs-r@mail.nih.gov
 Tissue Procurement: ATCC
 cDNA Library Preparation: Ling Hong/Rubin Laboratory
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
 Plate: LLC125 row: i column: 08
 High quality sequence stop: 467.
 Location/Qualifiers
 1. .649
 FEATURES
 source

/organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:3163855"
 /tissue_type="neuroblastoma"
 /lab_host="DH10B (phage-resistant)"
 /clone_lib="NIH_MGC_19"
 /note="Organ: brain; Vector: pOTB7; Site: 1: XhoI; Site 2:
 EcoRI; cDNA made by oligo-dT priming. Directionally
 cloned into EcoRI/XhoI sites using the following 5',
 adaptor: GGCACGAG(G). Library constructed by Ling Hong
 in the laboratory of Gerald M. Rubin (University of
 California, Berkeley) using ZAP-cDNA synthesis kit
 (Stratagene) and Superscript II RT (Life Technologies).
 Note: this is a NIH_MGC Library."
 BASE COUNT 93 a 250 c 212 g 94 t
 ORIGIN
 Query Match 19.1%; Score 334.8; DB 10; Length 649;
 Best Local Similarity 78.2%; Pred. No. 2.4e-25;
 Matches 482; Conservative 0; Mismatches 117; Indels 17; Gaps 6;
 QY 546 AGCGAGCGGACAGATCCAGTTGTGGCAGTTTCTACTGGAGCTGCTGGCAGCCGCGAA 605
 DB 3 AGAAAGCGGACAGATCCAGCTGTGGCAGTTTCTGCTGGAGCTGTGGCTGACCGCGAA 62
 QY 606 CGCCGCTGTCATCGCTGGGAGGGCGCCACGCGAGTTTCAAGCTCACGACCCCG-ACG 664
 DB 63 CGCCGCTGTCATCGCTGGGAGGGCGGTACGCGGAGTTTCAAGCTCACGACCCCGTAGC 122
 QY 665 AGTGCGGCGAGCTGGGGCGAGCGCAAGCAAGCCCAATATGAATACGACAAGCTAA 724
 DB 123 AGTGCGGCGCGGTGGGGCGAGCGCAAGCAAGCCCAATATGAATACGACAAGCTAA 182
 QY 725 GTGCGAGCTGCGCTACTACTAGCAAAACATCATGAGCAAGGTGCAATGCGCAAGCGCT 784
 DB 183 GCGCGCGCTTGCCTACTACTAGCAAAACATCATGAGCAAGGTGCAATGCGCAAGCGCT 242
 QY 785 AGCGCTTACCGCTTTCAGTTTCCAGGCGCTTGGCAGAGGCTTCCAGCCACCCCGCGCACG 844
 DB 243 AGCGCTTACCGCTTTCAGTTTCCAGGCGCTTGGCAGAGGCTTCCATGCGCGCGCGCGCA 302
 QY 845 CCCAGCGCGCGCTGCGCG-----CGCCGAGCGGCGAGCGCGCGCGCGCGCGCGCGCG 896
 DB 303 CGCTCAATGCGCGCGCGCGAGCTTGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 362
 QY 897 ACTTTACAGCTCCCGCTGGTCTGCTCCACTGCGCTTCCCGGCGCTTCCGAGCGCTTCCAA 956
 DB 363 TCTACGAAGCTGCGCGCGCGCGCTCGC-CCGCTGCGCTTCCCGCGCGCTTCTCCAACTCAA 421
 QY 957 CCTTATGCGAGCTGCGCGCGCGCTGGCGCGCTGCGCTTCTTACTGCGCTGGTCCC-- 1014
 DB 422 CCTCATGCGCGCTGCGCGCGGTGCGCGCGCGCTTCTCTACTTGGCGCGCGCTG 481
 QY 1015 --AAGCCACCGCGCTGCGCGCGCGCGCGCGCGCTTCTTCTTCTTCTTCTTCTTCTTCTT 1071
 DB 482 GCG 541
 QY 1072 CCCCCTCCCGCGCGCGCTTGGCGCGGTGGCGCGCTTCCACTTGGGGGGTCAATTATCAC 1131
 DB 542 CCCCCTCCCGCGCGCGCTTGGCGCGGTGGCGCGGTGGCGCGGTGGCGCGGTGGCGCGGT 599
 QY 1132 TAGACGGGACGCGCG 1147
 DB 600 TAGACGGGCGCGCG 615
 RESULT 13
 CB069127/c 451 bp mRNA linear EST 21-JAN-2003
 LOCUS is13c07.x1 HR85 islet Homo sapiens cDNA clone IMAGE:6364788 3',
 DEFINITION similar to TR:Q9581 Q9581 FEV PROTEIN.; mRNA sequence.
 ACCESSION CB069127

```

VERSION      CB069127.1  GI:27813647
KEYWORDS     EST.
SOURCE       Homo sapiens (human)
ORGANISM     Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE    1 (bases 1 to 451)
AUTHORS      Melton,D., Brown,J., Kenty,G., Permutt,A., Lee,C., Kaestner,K.,
              Lemishka,I., Scarce,M., Brestelli,J., Gradwohl,G., Clifton,S.,
              Hillier,L., Marra,M., Pape,D., Wyllie,T., Martin,J., Blistain,A.,
              Schmitt,A., Theising,B., Ritter,E., Ronko,I., Bennett,J., Cardenas
              M., Gibbons,M., McCann,R., Cole,R., Tsagareishvili,R., Williams,T.,
              Jackson,Y. and Bowers,Y.
              Endocrine Pancreas Consortium
              Unpublished
              Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue
              Endocrine Pancreas Consortium
              Harvard University, Howard Hughes Medical Institute
              Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,
              MA 02138
              Tel: 617-495-1812
              Fax: 617-495-8557
              Email: dmelton@biohp.harvard.edu
              Library was constructed by Dr. Hiroshi Inoue DNA sequencing by:
              Washington University Genome Sequencing Center For information on
              obtaining a clone please contact: Dr. Hiroshi Inoue
              (hinoue@im.wustl.edu)
              Seq primer: -40RP from Gibco
              High quality sequence stop: 439.

FEATURES     Location/Qualifiers
             1..451
                /organism="Homo sapiens"
                /mol_type="mRNA"
                /db_xref="taxon:9606"
                /clone="IMAGE:6364788"
                /tissue_type="Purified pancreatic islet"
                /lab_host="DH10B"
                /clone_lib="HR85 islet"
                /note="Organ: Pancreas; Vector: pBluescript SK(-); Site_1:
              NotI; Site_2: XhoI; cDNA made by oligo-dT priming.
              Size-selected on agarose gel. Average insert size ~1kb. 5'
              xhoI site was destroyed after directional cloning.
              Amplified once. Contact information: Hiroshi Inoue, MD,
              Metabolism Div. (Alan Permutt Lab), Washington University
              School of Medicine, Box 8127, 660 South Euclid Ave., St.
              Louis, MO 63110, E-mail: hinoue@imgate.wustl.edu, Tel:
              314-362-1916, Fax: 314-747-2692."

BASE COUNT   74 a 147 c 147 g 83 t
ORIGIN
Query Match      19.0%; Score 333.6; DB 14; Length 451;
Best Local Similarity 91.2%; Pred. No. 4.4e-25;
Matches 354; Conservative 0; Mismatches 34; Indels 0; Gaps 0;

QY 461 TGACCTACAGATCCGTCGGAGATGGTCTTTTAAAGGAAGCGAAGAGCCCGAGCTGGG 520
    |||
Db 389 TGTGCTTGAGATCCGTCGGAGACGGTCTTCAAGACGGGAGAACCCGAGCTGGG 330
    |||
QY 521 GGCCTGTAGCCCTCGGTACAGAAAGGAGCGGCGAGATCCAGTTGTGCGCAGTTTCTAC 580
    |||
Db 329 GGCCTGTAGCCCGCGGTTCAGAAAGGAGCGGCGAGATCCAGTGTGCGCAGTTTCTCG 270
    |||
QY 581 TGGAGCTGTGGCAGACCGCGCAACCGCGGTGTCATCCGTGGGAGGGCGGCACGGCG 640
    |||
Db 269 TGGAGCTGTGGCTGACCGCGCAACCGCGGTGTCATCCGTGGGAGGGCGGTACCGGGC 210
    |||
QY 641 AGTTCAAGCTCACCGACCCCGCAGAGTGGCGCGAGCTGGGCGGAGCGCAAGCAAGC 700
    |||
Db 209 AGTTCAAGCTCACCGACCCCGCAGAGTGGCGCGGCGGTGGCGGAGCGCAAGCAAGC 150
    |||
QY 701 CCATATGAACTACGACAGCTTAAGTCGAGCACTGCGCTACTACTACTACGACAAATCA 760
    |||
Db 149 CCAACATGAACACGAGCTGAGCGCGCCCTGCGCTACTACTACTACGACAAATCA 90
    |||

```

```

QY 761 TGAGCAAGGTGCAGCGCAAGCGCTACGCTACCGCTTTCACCTTCAGTTCGAGGCGCTGGCACAGG 820
    |||
Db 89 TGAGCAAGGTGCATGCAAGCGCTACGCTACCGCTTCGACTTCGACTTCGAGGCGCTGGCGCAGG 30
    |||
QY 821 CTTGCGCAGCCACCGCCGCGCAGGCCCA 848
    |||
Db 29 CTTGCGCAGCCGCGCGCGCGCGCAGGCTCA 2
    |||

RESULT 14
LOCUS    CB069363
DEFINITION is13c07.y1 HR85 islet Homo sapiens cDNA clone IMAGE:6364788 5'
          similar to TR:Q99581 Q99581 FEV PROTEIN. ; mRNA sequence.
ACCESSION CB069363
VERSION    CB069363.1 GI:27813883
KEYWORDS   EST.
SOURCE     Homo sapiens (human)
ORGANISM   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1 (bases 1 to 582)
AUTHORS      Melton,D., Brown,J., Kenty,G., Permutt,A., Lee,C., Kaestner,K.,
              Lemishka,I., Scarce,M., Brestelli,J., Gradwohl,G., Clifton,S.,
              Hillier,L., Marra,M., Pape,D., Wyllie,T., Martin,J., Blistain,A.,
              Schmitt,A., Theising,B., Ritter,E., Ronko,I., Bennett,J., Cardenas
              M., Gibbons,M., McCann,R., Cole,R., Tsagareishvili,R., Williams,T.,
              Jackson,Y. and Bowers,Y.
              Endocrine Pancreas Consortium
              Unpublished
              Other ESTs: is13c07.x1
              Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue
              Endocrine Pancreas Consortium
              Harvard University, Howard Hughes Medical Institute
              Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,
              MA 02138
              Tel: 617-495-1812
              Fax: 617-495-8557
              Email: dmelton@biohp.harvard.edu
              Library was constructed by Dr. Hiroshi Inoue DNA sequencing by:
              Washington University Genome Sequencing Center For information on
              obtaining a clone please contact: Dr. Hiroshi Inoue
              (hinoue@im.wustl.edu)
              Seq primer: -40UP from Gibco
              High quality sequence stop: 477.

FEATURES     Location/Qualifiers
             1..582
                /organism="Homo sapiens"
                /mol_type="mRNA"
                /db_xref="taxon:9606"
                /clone="IMAGE:6364788"
                /tissue_type="Purified pancreatic islet"
                /lab_host="DH10B"
                /clone_lib="HR85 islet"
                /note="Organ: Pancreas; Vector: pBluescript SK(-); Site_1:
              NotI; Site_2: XhoI; cDNA made by oligo-dT priming.
              Size-selected on agarose gel. Average insert size ~1kb. 5'
              xhoI site was destroyed after directional cloning.
              Amplified once. Contact information: Hiroshi Inoue, MD,
              Metabolism Div. (Alan Permutt Lab), Washington University
              School of Medicine, Box 8127, 660 South Euclid Ave., St.
              Louis, MO 63110, E-mail: hinoue@imgate.wustl.edu, Tel:
              314-362-1916, Fax: 314-747-2692."

BASE COUNT   98 a 195 c 205 g 84 t
ORIGIN
Query Match      18.8%; Score 329; DB 14; Length 582;
Best Local Similarity 90.9%; Pred. No. 1e-24;
Matches 350; Conservative 0; Mismatches 35; Indels 0; Gaps 0;

QY 461 TGACCTACAGATCCGTCGGAGATGGTCTTTTAAAGGAAGCGAAGAGCCCGAGCTGGG 520
    |||

```

```

Db      197  TGCTGCTTGAGATCCGCTGGAGACGGTCTTCTTCAAGGACGGGAAGAACCCGAGCTGGG 256
Qy      521  GGCCGCTGAGCCCTGGGTACAGAAAGGAGCGGCGAGATCCAGTTGTGCGAGTTTCTAC 580
      |||||
Db      257  GGCCGCTGAGCCCGCGGTTCAGAAAGGAGCGGACAGATCCAGCTGTGGCAGTTTCTGC 316
      |||||
Qy      581  TGGAGCTGCTGGCAGACCGCGGAGACCGCGGTGTCATCGCTGGAGGGCGGCCACGGCG 640
      |||||
Db      317  TGGAGCTGCTGGGTGACCGCGGACCGCGGTGTCATCGCTGGAGGGCGGTTCACGGCG 376
      |||||
Qy      641  AGTTCAAGCTCACGACCGGAGGTCGCGGAGCGCTGGGCGAGCGCCAGAGCAAGC 700
      |||||
Db      377  AGTTCAAGCTCACGACCGGAGGTCGCGGAGCGGTGGGCGGAGCGGAGAGCAAGC 436
      |||||
Qy      701  CCAATATGAATACGACGAACTAAGTCGAGCACTGCGCTACTACTACGACAAACATCA 760
      |||||
Db      437  CCAATATGAATACGACGAACTAAGTCGAGCACTGCGCTACTACTACTACGACAAACATCA 496
      |||||
Qy      761  TGAGCAAGGTGACGCGGAGCGCTAGCGCTTTCAGCTTTCAGGCGCTGGCACAGG 820
      |||||
Db      497  TGAGCAAGGTGACGCGGAGCGCTAGCGCTTTCAGCTTTCAGGCGCTGGCACAGG 556
      |||||
Qy      821  CTTGCCAGCCACACCGCGCACGC 845
      |||||
Db      557  GCTGCCAGCGCGCGCGCGCACGC 581

```

```

RESULT 15
BF470469
LOCUS
DEFINITION
  BF470469 458 bp mRNA linear EST 04-DEC-2000
  UI-M-BH3-avj-b-02-0-UI.r1 NIH_BMAP_M_S4 Mus musculus cDNA clone
  UI-M-BH3-avj-b-02-0-UI 5', mRNA sequence.
ACCESSION
  BF470469
VERSION
  BF470469.1 GI:11539652
KEYWORDS
  EST.
SOURCE
  Mus musculus (house mouse)
ORGANISM
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
  1 (bases 1 to 458)
  Bonaldo,M.F., Lennon,G. and Soares,M.B.
  Normalization and subtraction: two approaches to facilitate gene
  discovery
JOURNAL
  Genome Res. 6 (9), 791-806 (1996)
MEDLINE
  97044477
PUBMED
  8889548
COMMENT
  Contact: Chin, H
  National Institute of Mental Health
  6001 Executive Blvd. Room 7N-7190, MSC 9643, Bethesda, MD
  20892-9643, USA
  Tel: 301 443 1706
  Fax: 301 443 9890
  Email: mEST@mail.nih.gov
  cDNA Library Preparation: M.B. Soares Lab Clone distribution:
  Researchers may obtain BMAP cDNA clones from RESEARCH GENETICS. It
  should be noted that Bento Soares is generating a small number of
  additional specialized non-redundant arrays of BMAP cDNAs whose
  availability will be considered under appropriate and limited
  collaborative arrangements
  Seq primer: M13 Reverse.
  Location/Qualifiers
  1..458
    /organism="Mus musculus"
    /mol_type="mRNA"
    /strain="C57BL/6J"
    /db_xref="taxon:10090"
    /clone="UI-M-BH3-avj-b-02-0-UI"
    /dev_stage="27-32 days"
    /lab_host="DH10B (Life Technologies)"
    /clone_lib="NIH_BMAP_M_S4"
    /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
    polylinker. Site 1: Not I; Site 2: Eco RI; The
    NIH_BMAP_M_S4 library is a subtracted library of a series,

```

FEATURES

source

ultimately derived from a mixture of individually tagged normalized libraries from ten regions of the mouse brain (cerebellum, brain stems, olfactory bulbs, hypothalamus, cortex, amygdala, basal ganglia, pineal gland, striatum, hippocampus) after a series of subtractions to reduce the representation of cDNAs from which ESTs had already been generated. The following serially subtracted libraries were generated in this process: NIH_BMAP_M_S4, NIH_BMAP_M_S3.3, NIH_BMAP_M_S3.2, NIH_BMAP_M_S3.1, NIH_BMAP_M_S2, NIH_BMAP_M_S1. The subtracted library (NIH_BMAP_M_S4) was constructed as follows: PCR-amplified cDNA inserts from NIH_BMAP_M_S3.3, NIH_BMAP_M_S3.2, and NIH_BMAP_M_S3.1 clones from which 3' ESTs had been derived was used as a driver in a hybridization with a pool of the NIH_BMAP_M_S3.3, NIH_BMAP_M_S3.2, and NIH_BMAP_M_S3.1 libraries in the form of single-stranded circles. The remaining single-stranded circles (subtracted library) was purified by hydroxyapatite column chromatography, converted to double-stranded circles and electroporated into DH10B bacteria (Life Technologies) to generate the NIH_BMAP_M_S4 library. This procedure has been previously described (Bonaldo, Lennon and Soares, Genome Research 6:791-806, 1996)." 79 t

BASE COUNT
ORIGIN

Query Match 18.2%; Score 318; DB 10; Length 458;
Best Local Similarity 97.0%; Pred. NO. 1.7e-23;
Matches 324; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

```

Qy      470  CAGATCCCGTCGGAGATGTCCTTTTAAAGGAAGGAAGAGCCCGAGCTGGGGCCGCTGA 529
      |||||
Db      125  CAGATCCCGTCGGAGATGTCCTTTTAAAGGAAGGAAGAGCCCGAGCTGGGGCCGCTGA 184
      |||||
Qy      530  GCCCTGCGGTACAGAAAGCAGCGGCAGATCCAGTTGTGGCAGTTTCTCTGGAGCTGC 589
      |||||
Db      185  GCCCTGCGGTACAGAAAGCAGCGGCAGATCCAGTTGTGGCAGTTTCTCTGGAGCTGC 244
      |||||
Qy      590  TGGCAGACCGCGGAAACCGCGCTGCATCGCTGGGAGGGCGGCCACGGCGAGTTCAAGC 649
      |||||
Db      245  TGGCAGATCGCGGAAACCGCGCTTGCATCGCTGGGAGGGCGGCCACGGCGAGTTCAAGC 304
      |||||
Qy      650  TCACCGACCCCGACGAGGTGGCGCGACGCTGGGGCGAGCGCAAGAGCAAGCCCAATATGA 709
      |||||
Db      305  TCACCGACCCCGACGAGGTGGCGCGCTGGGGCGAGCGCAAGAGCAAGCCCAATATGA 364
      |||||
Qy      710  ACTAGACAGCTAAGTCGAGCACTGCGCTACTACTACGACAAAACATCATGAGCAAGG 769
      |||||
Db      365  ACTAGACAGCTAAGTCGCGGCTGCGCTACTACTACTACGACAAAACATCATGAGCAAGG 424
      |||||
Qy      770  TGCACGGCAAGCGCTACGCTACCGCTTTGACTT 803
      |||||
Db      425  TGCACGGCAAGCGCTACGCTACCGCTTTGACTT 458
      |||||

```

Search completed: November 26, 2003, 16:26:46
Job time : 3866 secs